

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2001, 13:22:11 ; Search time 18.61 Seconds  
(without alignments)  
1778.740 Million cell updates/sec

Title: US-09-689-430-2

Perfect score: 7869

Sequence: 1 M0E1STCFLLCRLRFCSA.....WVHQLRMVGLGCEAODLY (1471)

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID                  | Description        |
|------------|-------|-------------|--------|---------------------|--------------------|
| 1          | 7857  | 99.8        | 1471   | 1 US-08-683-839B-3  | Sequence 3, Appl1  |
| 2          | 7616  | 96.8        | 1661   | 2 US-08-882-083-2   | Sequence 2, Appl1  |
| 3          | 7616  | 96.8        | 1661   | 2 US-08-558-107-2   | Sequence 2, Appl1  |
| 4          | 7616  | 96.8        | 1661   | 4 US-09-243-539-2   | Sequence 2, Appl1  |
| 5          | 7419  | 94.3        | 2351   | 1 US-08-121-202-2   | Sequence 2, Appl1  |
| 6          | 7410  | 94.2        | 2351   | 6 5422260-1         | Patent No. 5422260 |
| 7          | 7406  | 94.1        | 2351   | 6 5171844-2         | Patent No. 5171844 |
| 8          | 7403  | 94.1        | 2351   | 1 US-08-366-851A-2  | Sequence 2, Appl1  |
| 9          | 7315  | 93.0        | 2332   | 1 US-07-864-004B-4  | Sequence 4, Appl1  |
| 10         | 7315  | 93.0        | 2332   | 1 US-08-251-937A-4  | Sequence 4, Appl1  |
| 11         | 7315  | 93.0        | 2332   | 1 US-08-212-133A-2  | Sequence 2, Appl1  |
| 12         | 7315  | 93.0        | 2332   | 1 US-08-474-503-2   | Sequence 2, Appl1  |
| 13         | 7315  | 93.0        | 2332   | 2 US-08-670-707A-2  | Sequence 2, Appl1  |
| 14         | 7315  | 93.0        | 2332   | 4 US-09-037-601-2   | Sequence 2, Appl1  |
| 15         | 7315  | 93.0        | 2332   | 5 PCT-US93-03275-4  | Sequence 4, Appl1  |
| 16         | 7315  | 93.0        | 2332   | 5 PCT-US94-13200-2  | Sequence 2, Appl1  |
| 17         | 7296  | 92.7        | 2332   | 1 US-08-276-594A-2  | Sequence 2, Appl1  |
| 18         | 7276  | 92.5        | 2332   | 4 US-09-324-867-3   | Sequence 3, Appl1  |
| 19         | 6552  | 83.3        | 1443   | 2 US-08-670-707A-39 | Sequence 39, Appl1 |
| 20         | 6552  | 83.3        | 1443   | 4 US-09-037-601-39  | Sequence 39, Appl1 |
| 21         | 6401  | 81.3        | 2343   | 4 US-09-324-867-2   | Sequence 2, Appl1  |
| 22         | 6347  | 80.7        | 2133   | 2 US-08-670-707A-37 | Sequence 37, Appl1 |
| 23         | 6347  | 80.7        | 2133   | 4 US-09-037-601-37  | Sequence 37, Appl1 |
| 24         | 6301  | 80.1        | 2319   | 1 US-08-212-133A-8  | Sequence 8, Appl1  |
| 25         | 6301  | 80.1        | 2319   | 1 US-08-474-503-6   | Sequence 6, Appl1  |
| 26         | 6301  | 80.1        | 2319   | 2 US-08-670-707A-6  | Sequence 6, Appl1  |
| 27         | 6301  | 80.1        | 2319   | 4 US-09-037-601-6   | Sequence 6, Appl1  |

|    |        |      |      |                     |                    |
|----|--------|------|------|---------------------|--------------------|
| 28 | 6301   | 80.1 | 2319 | 5 PCT-US94-13200-6  | Sequence 6, Appl1  |
| 29 | 6237.5 | 79.3 | 2115 | 4 US-09-324-867-5   | Sequence 5, Appl1  |
| 30 | 6196   | 78.7 | 2304 | 4 US-09-324-867-4   | Sequence 4, Appl1  |
| 31 | 2444.5 | 31.1 | 341  | 1 US-08-121-202-4   | Sequence 4, Appl1  |
| 32 | 2394.5 | 30.4 | 2183 | 3 US-08-746-111-5   | Sequence 5, Appl1  |
| 33 | 1726   | 21.9 | 368  | 1 US-08-212-133A-6  | Sequence 6, Appl1  |
| 34 | 1726   | 21.9 | 368  | 1 US-08-474-503-4   | Sequence 4, Appl1  |
| 35 | 1726   | 21.9 | 368  | 2 US-08-670-707A-4  | Sequence 4, Appl1  |
| 36 | 1726   | 21.9 | 368  | 4 US-09-037-601-4   | Sequence 4, Appl1  |
| 37 | 1726   | 21.9 | 368  | 5 PCT-US94-13200-4  | Sequence 4, Appl1  |
| 38 | 1722   | 21.9 | 367  | 1 US-07-864-004B-2  | Sequence 2, Appl1  |
| 39 | 1722   | 21.9 | 367  | 1 US-08-251-937A-2  | Sequence 2, Appl1  |
| 40 | 1722   | 21.9 | 367  | 5 PCT-US93-03275-2  | Sequence 2, Appl1  |
| 41 | 1153   | 14.7 | 218  | 1 US-07-607-538C-5  | Sequence 5, Appl1  |
| 42 | 1153   | 14.7 | 218  | 2 US-08-162-402B-5  | Sequence 5, Appl1  |
| 43 | 846    | 10.8 | 160  | 2 US-08-162-402B-15 | Sequence 15, Appl1 |
| 44 | 824    | 10.5 | 154  | 2 US-08-162-402B-17 | Sequence 17, Appl1 |
| 45 | 650    | 8.3  | 463  | 2 US-08-162-402B-9  | Sequence 9, Appl1  |

#### ALIGNMENTS

```
RESULT 1
US-08-683-839B-3
; Sequence, 3, Application US/08683839B
; Patent No. 5744326
; GENERAL INFORMATION:
; APPLICANT: ILL, Charles . R. et al.
; TITLE OF INVENTION: Use of Viral Cis-Acting Post-Transcriptional
; TITLE OF INVENTION: Regulatory Sequences To Increase Expression of
; TITLE OF INVENTION: Intronsless Genes Containing Near-Consensus Splice Sites
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,839B
; FILING DATE: 11-MARCH-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Remillard, Jane E.
; REGISTRATION NUMBER: 38,872
; REFERENCE/DOCKET NUMBER: TTI-138
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1471 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-683-839B-3

Query Match 99.8%; Score 7857; DB 1; Length 1471;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1469; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 M0E1STCFLLCRLRFCSATRRYVIGAVELSDMDYQSDLGELPVDAKRPVPRPKSPFN 60
```





Matches 1450; Conservative 5; Mismatches 10; Indels 202; Gaps 4;

```

OY 1 MOELSTCFELLCRECFATRRYYLGAVELSMDYMOGDLGELPYDARPPVPKSPFN 60
Db 1 MEELSTCFELLCRECFATRRYYLGAVELSMDYMOGDLGELPYDARPPVPKSPFN 60
OY 61 TSVVYKKTLEVEFTVHLEFNIAKPRPPMGLGPTIOAEVYDVITTLKNMASHPVSLNAV 120
Db 61 TSVVYKKTLEVEFTVHLEFNIAKPRPPMGLGPTIOAEVYDVITTLKNMASHPVSLNAV 120
OY 121 GVSYKASGAGYDQTSOREKEDKVPFGSGSHYVWOLKENGPMASDPLCLTYSLSH 180
Db 121 GVSYKASGAGYDQTSOREKEDKVPFGSGSHYVWOLKENGPMASDPLCLTYSLSH 180
OY 181 VDLVMDLNSGLIGALLVREGSLAKEKTOHLKFTLLFAVPDEGKSMSEFNKNSLMORD 240
Db 181 VDLVMDLNSGLIGALLVREGSLAKEKTOHLKFTLLFAVPDEGKSMSEFNKNSLMORD 240
OY 241 AASARAMPKMTVNGVYVRSJLPGJLIGCHRSYVNHVIGMGTTPVEHSTIFLEGHTELVNHN 300
Db 241 AASARAMPKMTVNGVYVRSJLPGJLIGCHRSYVNHVIGMGTTPVEHSTIFLEGHTELVNHN 300
OY 301 RASLEISPIFTFLTAOTLLMDLGOFLCHHSOHODGMEAVYKVDSCPEEPOLMKNE 360
Db 301 RASLEISPIFTFLTAOTLLMDLGOFLCHHSOHODGMEAVYKVDSCPEEPOLMKNE 360
OY 361 EAEYDDDLTDEMDVYFEEDDONSFIQISVAKKHPKTYWHTAAEEBMDAPLVA 420
Db 361 EAEYDDDLTDEMDVYFEEDDONSFIQISVAKKHPKTYWHTAAEEBMDAPLVA 420
OY 421 PDDRSYKQYLNNGPQIRGRYKVKVRYMAYDETFKTREALQIHESGILGPLLYGEVDTL 480
Db 421 PDDRSYKQYLNNGPQIRGRYKVKVRYMAYDETFKTREALQIHESGILGPLLYGEVDTL 480
OY 481 LIIFRNOASRPYNYPHGITTDRPLYSRRLKGVKHLKDFPLPEELIKRYWTVYVEDGP 540
Db 481 LIIFRNOASRPYNYPHGITTDRPLYSRRLKGVKHLKDFPLPEELIKRYWTVYVEDGP 540
OY 541 TKSDPRCLTRYSSFVNNERDLASGLJLPLICKESYVDGNOIMSKRWVILFSVDE 600
Db 541 TKSDPRCLTRYSSFVNNERDLASGLJLPLICKESYVDGNOIMSKRWVILFSVDE 600
OY 601 NRSWLTENIORFLPNPAGVLEDEPEQASNMHSINGVPSIQLSYCLHEVAYWYLS 660
Db 601 NRSWLTENIORFLPNPAGVLEDEPEQASNMHSINGVPSIQLSYCLHEVAYWYLS 660
OY 661 IGAQTDPLSVFFSGYTFKHKVYEDTLTLPFPGSETVFMSENGPGLWILGCHNSDFNRNG 720
Db 661 IGAQTDPLSVFFSGYTFKHKVYEDTLTLPFPGSETVFMSENGPGLWILGCHNSDFNRNG 720
OY 721 MTALKVSSC---DKNTGDY-----EDSYEDISAVLSKNNATPRFSONSRRPS 769
Db 721 MTALKVSSC---DKNTGDY-----EDSYEDISAVLSKNNATPRFSONSRRPS 769
OY 770 TROKOFNAT----- 778
Db 770 TROKOFNAT----- 778
OY 775 TROKOFNATIPENDIEKTDPMFAHRTPMKIQWVSSDMLMLRKOSTPHGLSLSDQE 834
Db 775 TROKOFNATIPENDIEKTDPMFAHRTPMKIQWVSSDMLMLRKOSTPHGLSLSDQE 834
OY 835 AKYTFSDDSPGALDSNNSLSEMTHFRPOLHSHGDMVFPPEGSLQJLRLNEKLGTTADPL 894
Db 835 AKYTFSDDSPGALDSNNSLSEMTHFRPOLHSHGDMVFPPEGSLQJLRLNEKLGTTADPL 894
OY 895 AMDNHGTQIKREEMKSOEKSPKTAFFKKDPTIISLNAACESNHAIAAINESQNKPELEVT 954
Db 895 AMDNHGTQIKREEMKSOEKSPKTAFFKKDPTIISLNAACESNHAIAAINESQNKPELEVT 954
OY 955 WAKOGRTERLCSQNPVYLKRHQREITRTTLOQDEIDYDTISVEKKEDFDIYDDEEN 1014
Db 955 WAKOGRTERLCSQNPVYLKRHQREITRTTLOQDEIDYDTISVEKKEDFDIYDDEEN 1014
OY 1015 OSPRSFQKTRHYFIAAVERLMDYGMSSSPVYLNRAOSGVSPOFKVYVFOEFTDGSFTQ 1074
Db 1015 OSPRSFQKTRHYFIAAVERLMDYGMSSSPVYLNRAOSGVSPOFKVYVFOEFTDGSFTQ 1074

```

```

OY 885 PLYRGELENEHLGILGPYIRAEVEDINWTFPRNOASRPYSVSSLTISYEDOROGAEPKRN 944
Db 1075 PLYRGELENEHLGILGPYIRAEVEDINWTFPRNOASRPYSVSSLTISYEDOROGAEPKRN 1134
OY 945 FVKPNETKYTFWKVOHNAAPTKEDEFCAMAYFSDVDEKDVHSGJLGPLLVCHTNTLNP 1004
Db 1135 FVKPNETKYTFWKVOHNAAPTKEDEFCAMAYFSDVDEKDVHSGJLGPLLVCHTNTLNP 1194
OY 1005 AHGROYVQOEALFETITDETKSWYFTENMEBNCRAPONIOEDPTFKENYRFAINCYI 1064
Db 1195 AHGROYVQOEALFETITDETKSWYFTENMEBNCRAPONIOEDPTFKENYRFAINCYI 1254
OY 1065 MDTLPLGLVMAQDQRIWRLWLLSMGSMENHTSHFSGHVFTVKKREYKNAALNLYPGVET 1124
Db 1255 MDTLPLGLVMAQDQRIWRLWLLSMGSMENHTSHFSGHVFTVKKREYKNAALNLYPGVET 1314
OY 1125 VEMLPKAGIRWRECLIGEHLAGKSTLELVYSNKCQPLGMAAGHIDFQITASGOYGO 1184
Db 1315 VEMLPKAGIRWRECLIGEHLAGKSTLELVYSNKCQPLGMAAGHIDFQITASGOYGO 1374
OY 1185 WAPKLARLHYSGSINAMSTKEPPSWIKVDLAPMIHIGIKTOGAROKFSSLYISOFTIMY 1244
Db 1375 WAPKLARLHYSGSINAMSTKEPPSWIKVDLAPMIHIGIKTOGAROKFSSLYISOFTIMY 1434
OY 1245 SLDGKKMOTYRGNSTGLMVEFGVNDSSGIRKHNIFNPITIRYIRLHPHTHSISTLME 1304
Db 1435 SLDGKKMOTYRGNSTGLMVEFGVNDSSGIRKHNIFNPITIRYIRLHPHTHSISTLME 1494
OY 1305 LMGCDLNSCMPPLGMEKSAISDAQITASSYPTNNEFATWSPSKARLHLOGRSNAMPQVNN 1364
Db 1495 LMGCDLNSCMPPLGMEKSAISDAQITASSYPTNNEFATWSPSKARLHLOGRSNAMPQVNN 1554
OY 1365 PKEMLOVDFOKTMKYVTGTOGVKSILTSMTYKFEPLISSODGHQWTLFONGKVKVYFOG 1424
Db 1555 PKEMLOVDFOKTMKYVTGTOGVKSILTSMTYKFEPLISSODGHQWTLFONGKVKVYFOG 1614
OY 1425 NODSFTPVVNSLDPLRLTRYLRHPQSWHOIALRMEVLGCEADLY 1471
Db 1615 NODSFTPVVNSLDPLRLTRYLRHPQSWHOIALRMEVLGCEADLY 1661

RESULT 3
US-08-558-107-2
; Sequence 2, Application US/08558107
; Patent No. 5910481
; GENERAL INFORMATION:
; APPLICANT: VOORBERG, Johannes J.
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/558,107
; FILING DATE: 13-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: ISACSON, John P.
; REGISTRATION NUMBER: 33,715
; REFERENCE/DOCKET NUMBER: 30472/212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399

```

TELEX: 904136  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1661 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-558-107-2

Query Match 96.8%; Score 7616; DB 2; Length 1661;  
 Best Local Similarity 87.0%; Pred. No. 0;  
 Matches 1450; Conservative 5; Mismatches 10; Indels 202; Gaps 4;

QY 1 MEIETSTCFCLLRFCSATRRYLGAVELSMQSDGLGELPVADARPPRPVPSFPPN 60  
 DB 1 MEIETSTCFCLLRFCSATRRYLGAVELSMQSDGLGELPVADARPPRPVPSFPPN 60  
 QY 61 TSVMYKKTLEVEFTVHLFNIAKPRPPMGLGPTIOAEVYDTVTITLKNMASHPVSLAHV 120  
 DB 61 TSVMYKKTLEVEFTVHLFNIAKPRPPMGLGPTIOAEVYDTVTITLKNMASHPVSLAHV 120  
 QY 121 GVSVMKASGAEYDQTSOREKEDKVPFGSGHTYVMVYLKENGPMASPLCLTYSLSH 180  
 DB 121 GVSVMKASGAEYDQTSOREKEDKVPFGSGHTYVMVYLKENGPMASPLCLTYSLSH 180  
 QY 121 GVSVMKASGAEYDQTSOREKEDKVPFGSGHTYVMVYLKENGPMASPLCLTYSLSH 180  
 DB 121 GVSVMKASGAEYDQTSOREKEDKVPFGSGHTYVMVYLKENGPMASPLCLTYSLSH 180  
 QY 181 VDLVMDLNSGLLGLALVREGSLAKEKQTTLKFTLLFAVPDEGKSWHSETKNSLMQDD 240  
 DB 181 VDLVMDLNSGLLGLALVREGSLAKEKQTTLKFTLLFAVPDEGKSWHSETKNSLMQDD 240  
 QY 181 VDLVMDLNSGLLGLALVREGSLAKEKQTTLKFTLLFAVPDEGKSWHSETKNSLMQDD 240  
 DB 181 VDLVMDLNSGLLGLALVREGSLAKEKQTTLKFTLLFAVPDEGKSWHSETKNSLMQDD 240  
 QY 241 AASARAPKMTVNGVYVNSRLSGLIGCHRRKSYVMHVGMTTPEVHSIFLEGHTELVNHH 300  
 DB 241 AASARAPKMTVNGVYVNSRLSGLIGCHRRKSYVMHVGMTTPEVHSIFLEGHTELVNHH 300  
 QY 301 RQASLEIPITFTLQTLMDLGOFLFCCHISSHQDGEAYKVDSCBEPQLRKNNH 360  
 DB 301 RQASLEIPITFTLQTLMDLGOFLFCCHISSHQDGEAYKVDSCBEPQLRKNNH 360  
 QY 301 RQASLEIPITFTLQTLMDLGOFLFCCHISSHQDGEAYKVDSCBEPQLRKNNH 360  
 DB 301 RQASLEIPITFTLQTLMDLGOFLFCCHISSHQDGEAYKVDSCBEPQLRKNNH 360  
 QY 361 EAEYDDDLTSEMVDVFRDDNSPSFIOIRSVAKKHPTWYHYIAEEDMDYAPLVIA 420  
 DB 361 EAEYDDDLTSEMVDVFRDDNSPSFIOIRSVAKKHPTWYHYIAEEDMDYAPLVIA 420  
 QY 421 PDDRSYKQYLNNGPQIRGRKKYVFMAATDEFTKREAIQHEGSLGLPLLYGEVDTL 480  
 DB 421 PDDRSYKQYLNNGPQIRGRKKYVFMAATDEFTKREAIQHEGSLGLPLLYGEVDTL 480  
 QY 481 LIIRKNOASRPYNTYPHGTTDVRPLXSRRLPKGVKHLKDFPLPGELFFKMTVTVEDGP 540  
 DB 481 LIIRKNOASRPYNTYPHGTTDVRPLXSRRLPKGVKHLKDFPLPGELFFKMTVTVEDGP 540  
 QY 541 TKSDPRCLTRYYSFVNNERDLASGLIPLLICYESVDQKNGQIMSDKRNVLFSVDE 600  
 DB 541 TKSDPRCLTRYYSFVNNERDLASGLIPLLICYESVDQKNGQIMSDKRNVLFSVDE 600  
 QY 601 NRSWLTENIORFLPNPAGVQLEDEPQASINIMHSINGVEDSLQVCLHEVAYWYIIS 660  
 DB 601 NRSWLTENIORFLPNPAGVQLEDEPQASINIMHSINGVEDSLQVCLHEVAYWYIIS 660  
 QY 661 IGAQTDLSVFEFSQYTFKHKMYEDTLTLPFSGETVFMSPENPGIMLGCNHSDRNNG 720  
 DB 661 IGAQTDLSVFEFSQYTFKHKMYEDTLTLPFSGETVFMSPENPGIMLGCNHSDRNNG 720  
 QY 721 MTALLKVVSSCIPGEEEDDYLDLEKIFSEDYDI-----VDSLIFRFSQNSRHS 774  
 DB 721 MTALLKVVSSCIPGEEEDDYLDLEKIFSEDYDI-----VDSLIFRFSQNSRHS 774  
 QY 770 TRQKOFNAT----- 778  
 DB 770 TRQKOFNAT----- 778  
 QY 775 TRQKOFNATITPENDIEKTDPMFAHRTMPKIQNVSSDMLMLRQSPPHGLSLDLOE 834  
 DB 775 TRQKOFNATITPENDIEKTDPMFAHRTMPKIQNVSSDMLMLRQSPPHGLSLDLOE 834  
 QY 779 ----- 778  
 DB 779 ----- 778  
 QY 835 AKYETFSDDPSGALDSNNSISEMTHFRPOLHHSQDMVFTPESGLQLRLNEKLGTTADPL 894  
 DB 835 AKYETFSDDPSGALDSNNSISEMTHFRPOLHHSQDMVFTPESGLQLRLNEKLGTTADPL 894

QY 779 ----- 778  
 DB 895 AMDNHYGTQIPKEEKKSOEKSPKTAFAKKKDTILSLNACESNHAIAINEGONKPEIEVT 954  
 QY 779 -----PPYLKRHOREITRTTTLQSDOEIEDYDOTTISYEMKKEPDIYEDEN 824  
 DB 955 MAKOGTERLCSQNPVPLKRHOREITRTTTLQSDOEIEDYDOTTISYEMKKEPDIYEDEN 1014  
 QY 825 OSPRSFOKKTRHYFTAAYVERLMDYGMSSSPHYLRNRAOGSGVPOKKVYVFOETDGSFTO 884  
 DB 1015 OSPRSFOKKTRHYFTAAYVERLMDYGMSSSPHYLRNRAOGSGVPOKKVYVFOETDGSFTO 1074  
 QY 885 PLYRGLNELHGLLGPYIRAEVEDNIMYTFRNQASRPYSFYSLSLISYEDDROGAEPKRN 944  
 DB 1075 PLYRGLNELHGLLGPYIRAEVEDNIMYTFRNQASRPYSFYSLSLISYEDDROGAEPKRN 1134  
 QY 945 FYKPNETKYTFKKVQHNAPYKDEEDCKAMAYFSVDLEKDYHSLIGPLLVCHTNLNP 1004  
 DB 1135 FYKPNETKYTFKKVQHNAPYKDEEDCKAMAYFSVDLEKDYHSLIGPLLVCHTNLNP 1194  
 QY 1005 AHGRQVTOEFALEFTIPEDETSWYFTENMERNCRAPCNIOMEPTPEKENTRPHANGYI 1064  
 DB 1195 AHGRQVTOEFALEFTIPEDETSWYFTENMERNCRAPCNIOMEPTPEKENTRPHANGYI 1254  
 QY 1065 MDTLPLVMAODORTIRWYLLSMGSENENHSIHFSGHVFTVRKKEEYKALYNLYPGVEET 1124  
 DB 1255 MDTLPLVMAODORTIRWYLLSMGSENENHSIHFSGHVFTVRKKEEYKALYNLYPGVEET 1314  
 QY 1125 VEMLPKAGIWRBECLIGELHLAGMSTPLFYVSNKQPTPLCMASGHIRDPOITRAGQYQ 1184  
 DB 1315 VEMLPKAGIWRBECLIGELHLAGMSTPLFYVSNKQPTPLCMASGHIRDPOITRAGQYQ 1374  
 QY 1185 NAKPLARLHYSGSINAMSKPEFSMLKYVDLAPMTHIKIKOGAROKSSLYISQITMY 1244  
 DB 1375 NAKPLARLHYSGSINAMSKPEFSMLKYVDLAPMTHIKIKOGAROKSSLYISQITMY 1434  
 QY 1245 SLDGKKWQYTRNGSTGTTLVWFGVNDSSGIKHINIPPIIARYIRLHPTHYRSINSTLME 1304  
 DB 1435 SLDGKKWQYTRNGSTGTTLVWFGVNDSSGIKHINIPPIIARYIRLHPTHYRSINSTLME 1494  
 QY 1305 LMGCDLNSCMLPGMESKASIDAOITASSYFTNMATSPSKARLHLQGRNAMPVYNN 1364  
 DB 1495 LMGCDLNSCMLPGMESKASIDAOITASSYFTNMATSPSKARLHLQGRNAMPVYNN 1554  
 QY 1365 PKEMLOVDFOKTMKYTGVTGCVKSLTSMYKEFLISSODGHOWTLFPONGKVKYVQ 1424  
 DB 1555 PKEMLOVDFOKTMKYTGVTGCVKSLTSMYKEFLISSODGHOWTLFPONGKVKYVQ 1614  
 QY 1425 NQDSFTPVVNSLDPPLTFRYLRIHPQSWVHOIALRMEVYLGCEAODLY 1471  
 DB 1615 NQDSFTPVVNSLDPPLTFRYLRIHPQSWVHOIALRMEVYLGCEAODLY 1661

RESULT 4  
 US-09-243-539-2  
 Sequence 2, Application US/09243539  
 Patent No. 6130203  
 GENERAL INFORMATION:  
 APPLICANT: VOORBERG, Johannes J.  
 TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY  
 NUMBER OF SEQUENCES: 17  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Foley & Lardner  
 STREET: 3000 K Street, N.W., Suite 500  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20007-5109  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS





QY 779 ----- 778  
Db 1561 GKVPFLRVATESSAKPTSKLDBPLAMDNHYGTQIPKEEWSOEKSPKTAFFKKTDTLTL 1620  
QY 779 ----- -PVLAKRIQRETTTTLQSDOE 800  
Db 1621 MACESNHAIAINEGONKEIEVTYAKOGRTERLCSQNPVPLKRHOEITRTTLQSDOE 1680  
QY 801 IDYDDITISYEMKKEPDIDEDENOSPRSFOKTRHYFIAAVERLMDYMSSSPHYLRNR 860  
Db 1681 IDYDDITISYEMKKEPDIDEDENOSPRSFOKTRHYFIAAVERLMDYMSSSPHYLRNR 1740  
QY 861 AOSGSVPOKRVKVFQOEFTDGSFTOPLYRGELNEHLGLGPLYRAEVEDNINVTFRNOASR 920  
Db 1741 AOSGSVPOKRVKVFQOEFTDGSFTOPLYRGELNEHLGLGPLYRAEVEDNINVTFRNOASR 1800  
QY 921 PVSFYSLSLSTYEDDOROGAPRKNFVKNMETKTYFWKVOHNAAPTKDEPDCRAMAYFSV 980  
Db 1801 PVSFYSLSLSTYEDDOROGAPRKNFVKNMETKTYFWKVOHNAAPTKDEPDCRAMAYFSV 1860  
QY 981 DLEKDVHSLIGPLLYCHNTNTPAHGROYVQOEALFETIDETKSNVETENMERNCRA 1040  
Db 1861 DLEKDVHSLIGPLLYCHNTNTPAHGROYVQOEALFETIDETKSNVETENMERNCRA 1920  
QY 1041 PCNIOMEDPFEKENYRFHAINCYIMDTPLGLVMAODQIRIWMYLLSGSNENHSHIFSGH 1100  
Db 1921 PCNIOMEDPFEKENYRFHAINCYIMDTPLGLVMAODQIRIWMYLLSGSNENHSHIFSGH 1980  
QY 1101 VETVRRKEEYKALNLYPGVEFYEYEMLPKAGIMRECLTGEHLHAGNSTFLVYSNKC 1160  
Db 1981 VETVRRKEEYKALNLYPGVEFYEYEMLPKAGIMRECLTGEHLHAGNSTFLVYSNKC 2040  
QY 1161 QPPLGASGHINDFOITAGOGVONAPKLARLHSGSINAMSTKEPESIKYDLDAPMTI 1220  
Db 2041 QPPLGASGHINDFOITAGOGVONAPKLARLHSGSINAMSTKEPESIKYDLDAPMTI 2100  
QY 1221 HGKTOGAKOFKSSLYISQFIIMYSLDGKKMOTYRGNSTGLTLMVFEGVNDSSGKHNIFN 1280  
Db 2101 HGKTOGAKOFKSSLYISQFIIMYSLDGKKMOTYRGNSTGLTLMVFEGVNDSSGKHNIFN 2160  
QY 1281 PPIIARIYLIHPTHYRSINSTLMELMGCDLNSCSPMLGMSKASISDAOITASSYFTNMFA 1340  
Db 2161 PPIIARIYLIHPTHYRSINSTLMELMGCDLNSCSPMLGMSKASISDAOITASSYFTNMFA 2220  
QY 1341 TWSPSKARLHLDGRSNAMPPOVNNPKEMLOVDFOKTMKVTGVTGQVSLTSMVKEFL 1400  
Db 2221 TWSPSKARLHLDGRSNAMPPOVNNPKEMLOVDFOKTMKVTGVTGQVSLTSMVKEFL 2280  
QY 1401 ISSSQDGHOWTLFQNGKVKYFQGNODSFTPYVNSLIDPLRLTRYLRIRHPOSVWQIALRM 1460  
Db 2281 ISSSQDGHOWTLFQNGKVKYFQGNODSFTPYVNSLIDPLRLTRYLRIRHPOSVWQIALRM 2340  
QY 1461 EYLGCEADOLY 1471  
Db 2341 EYLGCEADOLY 2351

; FILLING DATE: 18-NOV-1986  
; APPLICATION NUMBER: 868,410  
; FILLING DATE: 29-MAY-1986  
; SEQ ID NO:1  
5422260-1  
LENGTH: 2351  
Query Match 94.2%; Score 7410; DB 6; Length 2351;  
Best Local Similarity 62.5%; Pred No. 0;  
Matches 1470; Conservative 0; Mismatches 1; Indels 880; Gaps 1;  
QY 1 MOELSTCFELCLNRCFSATRRYYLAGVELSMDYMQSDGLGELPYDAREPPVPKSFSPN 60  
Db 1 MOELSTCFELCLNRCFSATRRYYLAGVELSMDYMQSDGLGELPYDAREPPVPKSFSPN 60  
QY 61 TSVVYKKTLEFVETVHLFNIAKPRPPMGLGPTIOAEYDVTIVTTLKMAHSPVSLNAV 120  
Db 61 TSVVYKKTLEFVETVHLFNIAKPRPPMGLGPTIOAEYDVTIVTTLKMAHSPVSLNAV 120  
QY 121 GVSYMKASGAEYDQTSOREKEDDKVPFGSGHTYVMOYLKENGPMASDPLCLTYSLSH 180  
Db 121 GVSYMKASGAEYDQTSOREKEDDKVPFGSGHTYVMOYLKENGPMASDPLCLTYSLSH 180  
QY 181 VDLVRDLNSGLIGALLVCREGSLAKEKTQTLHKFTLLFAVFDEGKSWHSETKNSLMODRD 240  
Db 181 VDLVRDLNSGLIGALLVCREGSLAKEKTQTLHKFTLLFAVFDEGKSWHSETKNSLMODRD 240  
QY 241 AASARAMPMTNVNGVYVNSLPLGLIGCHRSYVWHVIGMGTTPVHSHFLBCHTLVNRH 300  
Db 241 AASARAMPMTNVNGVYVNSLPLGLIGCHRSYVWHVIGMGTTPVHSHFLBCHTLVNRH 300  
QY 301 ROSLEISPIFTLTAOTLLMDLGOFLFCCHLSHQHDGEAVYKVYDSCPEEPQLMKKNE 360  
Db 301 ROSLEISPIFTLTAOTLLMDLGOFLFCCHLSHQHDGEAVYKVYDSCPEEPQLMKKNE 360  
QY 361 EAEDYDDDLTJSEMDEVFRDDNSPFIQIRSVAKHRTWVHYTAAEEDMDVAPLYLA 420  
Db 361 EAEDYDDDLTJSEMDEVFRDDNSPFIQIRSVAKHRTWVHYTAAEEDMDVAPLYLA 420  
QY 421 PDDRSYKSOYLLNNGQRIQIRKRYKAVFMAAYTDEFTKTRBAIOHESGILGPLYGEGDTL 480  
Db 421 PDDRSYKSOYLLNNGQRIQIRKRYKAVFMAAYTDEFTKTRBAIOHESGILGPLYGEGDTL 480  
QY 481 LIIFKNOASRPNIYPHGTTDVRPLYSRRLPKGVNHLMDPFLPBEIRKYKMTVVEGP 540  
Db 481 LIIFKNOASRPNIYPHGTTDVRPLYSRRLPKGVNHLMDPFLPBEIRKYKMTVVEGP 540  
QY 541 TKS DPRCLTRYYSFVNNERDLASGLIGPLCYKESYDORQNOJMSDKRNVILFSVDE 600  
Db 541 TKS DPRCLTRYYSFVNNERDLASGLIGPLCYKESYDORQNOJMSDKRNVILFSVDE 600  
QY 601 NRSWLTENIQOFLRNPAGVQLEDEPFOASNIMHSINGVYDSDLSVCLHEVAWYLLS 660  
Db 601 NRSWLTENIQOFLRNPAGVQLEDEPFOASNIMHSINGVYDSDLSVCLHEVAWYLLS 660  
QY 661 IGAOTDFLSVFPFGYTFKHMYEDTLFLFPFGSETVPMSPENPGLWILGCHNSFRNRG 720  
Db 661 IGAOTDFLSVFPFGYTFKHMYEDTLFLFPFGSETVPMSPENPGLWILGCHNSFRNRG 720  
QY 721 MTALLKVSCKDKNTDYEDSYEDISAVLLSKNNAIEPRFSQNSRHPSTROKOPNAT-- 778  
Db 721 MTALLKVSCKDKNTDYEDSYEDISAVLLSKNNAIEPRFSQNSRHPSTROKOPNAT-- 778  
QY 779 ----- 778  
Db 781 PENDIEKTDPMFAHRTPMKIQONVSSDMLMLROSPTPHGLSLSDLOEAKYTFSDPS 840  
QY 779 ----- 778  
Db 841 PGADISNNSLSMTYFRQLHHSQDMVTPESGDLRLNEKLGITTAATLELKKLDFKVSST 900  
QY 779 ----- 778

```

Db 901 SNNLISTIPSDNLAAGTNTSSLGPPSMPVHYDSQDITLLEGGKSSPLESGLSLEE 960
Qy 779 ----- 778
Db 961 NNDKLESLGSLNDSQSSGKNVSTESGRLEFKRAHGRALLTKONALEKVISILKTN 1020
Qy 779 ----- 778
Db 1021 KTSNNSATNRKTHIDGPSLIIENSPSWONILESDTEFKKVTPLIHDRMLDKNATLRL 1080
Qy 779 ----- 778
Db 1081 NMSKNTTSSKNMENVQOKKEGP IPPDAONPDMSFFKMLFLPESARWIOETHGKNSLNSG 1140
Qy 779 ----- 778
Db 1141 QGSPKQVLGSLGPEKSVESQGNLSEKNKYVGGEFTKDVGLKEMVFPSSRNLELTINLDN 1200
Qy 779 ----- 778
Db 1201 LHENNTHNOEKKIOEIEKKEKTLIOENVLPLQIHVTGTKNEMKNLFLSTRQNVESYE 1260
Qy 779 ----- 778
Db 1261 GAYAPVLDFRSLNDSTNFTKKAHTAHFSKKEEENLEGLGNQTKOIVERYACTTRISPMT 1320
Qy 779 ----- 778
Db 1321 SQONVTVQSRKALQOFRLPLETELEKRIYDDTSTQMSKNKHLTPSTLQIDYNEKE 1380
Qy 779 ----- 778
Db 1381 KGAITOSPLSDCLTRSHSIPQANRSPPIAKVSPFSIRPIYLRVLFODNSHLPAASY 1440
Qy 779 ----- 778
Db 1441 RKKDSGOBSSHFLQAGAKKNLNLALITLMTGDOREVSLGTSATNSVTKKVENTVLP 1500
Qy 779 ----- 778
Db 1501 KPDLPTSGKVELLRKVNHYQKDLPTETNSGPHGLDLVEGSLLOCTGCAIKMNEANRP 1560
Qy 779 ----- 778
Db 1561 GKVPFLRVATESSAKTPSKLLDPLAMDNHYGTQIPKEEMKSOEKSEPTAFKKDITLSL 1620
Qy 779 ----- 778
Db 1621 NACESNHAIAINEGONKPELEVYAKOGRTERLCSQNPVPLKRHOQRETRITTLQSDOE 1680
Qy 779 ----- 778
Db 1681 IDYDDTISVEMKEDFDIYDEDENSPRSFQKTRHYFAAVERLMDYGMSSSPHYLRNR 860
Db 1681 IDYDDTISVEMKEDFDIYDEDENSPRSFQKTRHYFAAVERLMDYGMSSSPHYLRNR 1740
Qy 861 AOSGVPQPKVYFOEFTDGSFTQPLYRGELNENHGLGPGYRAEVEDNIWTFRNOQR 920
Db 1741 AOSGVPQPKVYFOEFTDGSFTQPLYRGELNENHGLGPGYRAEVEDNIWTFRNOQR 1800
Qy 921 PYSFSSSLISYEEDROGAEPRKNFVKPNETKTYEMKYOHNAHAPRTDEEDCKAMAEFSV 980
Db 1801 PYSFSSSLISYEEDROGAEPRKNFVKPNETKTYEMKYOHNAHAPRTDEEDCKAMAEFSV 1860
Qy 981 DLEKDVHSGSLGPLLVCHTNTLNPAGROVYQEFALFETIFDETQSWYFTENMERNCRA 1040
Db 1861 DLEKDVHSGSLGPLLVCHTNTLNPAGROVYQEFALFETIFDETQSWYFTENMERNCRA 1920
Qy 1041 PCNTIOMEDPTEKENTRFAINQYIMDTPLGLVYADQRIRWYLLSGNSNENHSHISFSGH 1100
Db 1921 PCNTIOMEDPTEKENTRFAINQYIMDTPLGLVYADQRIRWYLLSGNSNENHSHISFSGH 1980
Qy 1101 VFTYRKKKEEYKALYNLYPGVETVEMLPKSGKIRVELLIGEHLAGNSTLFLVYSNKC 1160

```

```

Db 1981 VFTYRKKKEEYKALYNLYPGVETVEMLPKSGKIRVELLIGEHLAGNSTLFLVYSNKC 2040
Qy 1161 QPPLGASGHIRDFQITASGOYQOMAPKELARLHYSGSINAMSTKEPFMIKVDLAPMI 1220
Db 2041 QPPLGASGHIRDFQITASGOYQOMAPKELARLHYSGSINAMSTKEPFMIKVDLAPMI 2100
Qy 1221 HGKTOGAKOKSSSLYSOPIIMYSLDGKKWOTYRGNSTGLMVFPGVNDSSGKIHNIEN 1280
Db 2101 HGKTOGAKOKSSSLYSOPIIMYSLDGKKWOTYRGNSTGLMVFPGVNDSSGKIHNIEN 2160
Qy 1281 PPIIARYILPHPHYSIRSTLMELMGCDLNSCSPMLGSKAISDAOITASSYFTNMEA 1340
Db 2161 PPIIARYILPHPHYSIRSTLMELMGCDLNSCSPMLGSKAISDAOITASSYFTNMEA 2220
Qy 1341 TWSPSKARLHLOGRSNAMPQVNNPKEMLOVDFOKTMKYGTGTQGVSLTSMYKEFL 1400
Db 2221 TWSPSKARLHLOGRSNAMPQVNNPKEMLOVDFOKTMKYGTGTQGVSLTSMYKEFL 2280
Qy 1401 ISSODGHOWTLFPONGKVKVFOGNDSFTPVVNSLDPELRLRYLRIRHPSWVQIALRM 1460
Db 2281 ISSODGHOWTLFPONGKVKVFOGNDSFTPVVNSLDPELRLRYLRIRHPSWVQIALRM 2340
Qy 1461 EYLGCBAODLY 1471
Db 2341 EYLGCBAODLY 2351

```

```

RESULT 7
5171844-2
; Patent No. 5171844
; APPLICANT: VAN OOEYEN, ALBERT J.J.; PANNEKOEK, HANS; VERBEET,
; MARTINUS P.; VAN LEEUW, ROBERT W.
; TITLE OF INVENTION: PROTEINS WITH FACTOR VIII ACTIVITY
; PROCESS FOR THEIR PREPARATION USING GENETICALLY-ENGINEERED CELLS
; AND PHARMACEUTICAL COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/205,226
; FILING DATE: 10-JUN-1988
; SEQ ID NO: 2:
; LENGTH: 2351
5171844-2

```

```

Query Match 94.1%; Score 7406; DB 6; Length 2351;
Best Local Similarity 62.5%; Pred. No. 0;
Matches 1469; Conservative 0; Mismatches 2; Indels 880; Gaps 1;

Qy 1 MOELSTCFPLCLLRFCGSAATRRYLLGAVELSDWYQSDGLBEPVDAFPPRVKSPFPN 60
Db 1 MOELSTCFPLCLLRFCGSAATRRYLLGAVELSDWYQSDGLBEPVDAFPPRVKSPFPN 60
Qy 61 TSVYVKKTLFEFTYHLFPIAKPRPPMGLGPTQAEVYDVTYTLKNMASHPVSLHAV 120
Db 61 TSVYVKKTLFEFTYHLFPIAKPRPPMGLGPTQAEVYDVTYTLKNMASHPVSLHAV 120
Qy 121 GVSYKASGAEYDDOTSQREKEDKVPFGSGHTYVMQVLKENGFMASDPLCLTYSLSH 180
Db 121 GVSYKASGAEYDDOTSQREKEDKVPFGSGHTYVMQVLKENGFMASDPLCLTYSLSH 180
Qy 181 VDLVYDNLNGLIGALLVCEGSLAKEKQTLHKFLLFAVDEGKSWSEFRKNSLMODRD 240
Db 181 VDLVYDNLNGLIGALLVCEGSLAKEKQTLHKFLLFAVDEGKSWSEFRKNSLMODRD 240
Qy 241 AASARAMPKMHVNVYNSLPLGICGRKRSYVMHVIGMGTPEVHSFELGHTFIVRNH 300
Db 241 AASARAMPKMHVNVYNSLPLGICGRKRSYVMHVIGMGTPEVHSFELGHTFIVRNH 300
Qy 301 ROASLEISPTFLTAQTLLMDGQFLFCCHISSHQHDMEAAYVVDSCPEEPQRLRMKNE 360
Db 301 ROASLEISPTFLTAQTLLMDGQFLFCCHISSHQHDMEAAYVVDSCPEEPQRLRMKNE 360
Qy 361 EAEDYDDDLTDEMDVVRFDNDNSPSFTQIRSVAKKHPKRWVHYIAAEEDMDYAPVLA 420

```

```

|||||
Db 361 EADYDDDLTJDESEMDYVRFEDDNDSPFIOIRSVAKKHPKTHVHYIAAEEDMDYAPLVLA 420
Qy 421 PDORSKSOYLNGNPGORIGRKKYKVFMAVYDETFKTRREALIOHESGIIIGPLLYGVGDTL 480
Db 421 PDORSKSOYLNGNPGORIGRKKYKVFMAVYDETFKTRREALIOHESGIIIGPLLYGVGDTL 480
Qy 481 LIIFFNOASRPNIYPHGIIIDVRPLYSRRLPKGVHLLKDFILPEIEIFKKYKTVVEDGP 540
Db 481 LIIFFNOASRPNIYPHGIIIDVRPLYSRRLPKGVHLLKDFILPEIEIFKKYKTVVEDGP 540
Qy 541 TKSDFRCLTRYSSFFVNMRDLASGLIGPLLYCYESVDQRCNOJMSDKRNVILFVDE 600
Db 541 TKSDFRCLTRYSSFFVNMRDLASGLIGPLLYCYESVDQRCNOJMSDKRNVILFVDE 600
Qy 601 NNSWLTENIOFELPNAGVYOLEDEPEFQASNTMHSINGVDSLOLSCHEVAWYILS 660
Db 601 NNSWLTENIOFELPNAGVYOLEDEPEFQASNTMHSINGVDSLOLSCHEVAWYILS 660
Qy 661 IGAOTDFLSVFSSGYTFKHKMYEDTLFLFPSSGTFVMSMNPGLMILGCHNSDFRRNG 720
Db 661 IGAOTDFLSVFSSGYTFKHKMYEDTLFLFPSSGTFVMSMNPGLMILGCHNSDFRRNG 720
Qy 721 MTALLKVVSCOKNTGDYEDSYEDISAYLLSKNNAIEPRFSQNSRHSSTROKOFNAT - 778
Db 721 MTALLKVVSCOKNTGDYEDSYEDISAYLLSKNNAIEPRFSQNSRHSSTROKOFNAT - 780
Qy 779 ----- 778
Db 781 PENDIEKTDPMFAHRTPMFKIONVSSDDLMLLROSPTPHGISLSLOEAKYETFSDDPS 840
Qy 779 ----- 778
Db 841 PGALDSNNSLSMTHFRPOLHHSQDMVTFPESGLOLRINEKIGTTATELKKIDFKVST 900
Qy 779 ----- 778
Db 901 SNMLISTIPDNLAAGTNTSSIGPPSMVHYDSQDLOTLFGKSSPLTESGGLSLSEE 960
Qy 779 ----- 778
Db 961 NNDKSLLESGLMNSQESSMGKNVSSTESGRLFKGRAGPALLTOKNALFKVISISLKTN 1020
Qy 779 ----- 778
Db 1021 KTSNNSATNRKTHIDGPSELLIENSPVONILIESDTEFKKVTPLIHDMLDKNATLRL 1080
Qy 779 ----- 778
Db 1081 NHMSNKTSSKNMEMVQOKKEGPPIPPAQNPDMSFFKMLFLPESARWIOPTHGKNSLSNG 1140
Qy 779 ----- 778
Db 1141 QGSPKQVLVLGPEKSVBQONFLSEKNKVYVKGFEFTKDVGLKEMVFPSSRMLFLTNLDN 1200
Qy 779 ----- 778
Db 1201 LHENNTHOEKKIOEIEIKKETLIOENVVLOIHTVTGKNFMKMLFLSTRONVGSYD 1260
Qy 779 ----- 778
Db 1261 GAYAPVLQDFRSLNDSTNRKTHAHFSKGBEENLBSLGNOTKOIVEKACTTRISPNT 1320
Qy 779 ----- 778
Db 1321 SQONFVTQSRKALKOFLPLEETELEKRIIVDSTQMSKNMKHLNPSLTLOIDYNEKE 1380
Qy 779 ----- 778
Db 1381 KGATIOSPLSDCLTRSHSIIPOANKSPFLIAKVSSPSTIRPIYLTIVLFDONNSHLPAASY 1440
Qy 779 ----- 778

```

```

Db 1441 RKKDSGVQESSHFILOGAKNNLSAILTLEMTGDQREVSLGTSATNSVTKKVENTVLP 1500
Qy 779 ----- 778
Db 1501 KPDLFKTSGKVELLPKHVHIYOKDLFPTETSNGPSGHLDLVBGSLQCTEGAIKNNEANRP 1560
Qy 779 ----- 778
Db 1561 GKVPFLRATASSAKTPSKLLDPLAMDNHYGTQIPKEEMKSOEKSPEKTAFKKKDTIISL 1620
Qy 779 ----- 778
Db 1621 NACESNHAIAINSONKPEIEVTAQOKRTERLCQNDPVLKRRHOREITRTTLOSDOE 1680
Qy 801 IDYDITISEVEMKKEDEFDIYDEDENSPSPQKTPHYTIAAVERIMDYGSSPVLNRR 860
Db 1681 IDYDITISEVEMKKEDEFDIYDEDENSPSPQKTPHYTIAAVERIMDYGSSPVLNRR 1740
Qy 861 AOSGSVPQFKKVFQEFIDGFTQPLRGELNEHLGLGPYIRAEVEDNIMVTFPRNOASR 920
Db 1741 AOSGSVPQFKKVFQEFIDGFTQPLRGELNEHLGLGPYIRAEVEDNIMVTFPRNOASR 1800
Qy 921 PYSFYSLSIYEEEDQROGAEPKRNKFNKNEKTYFKVQVHMAPTKDEFDCKAMAYFSDV 980
Db 1801 PYSFYSLSIYEEEDQROGAEPKRNKFNKNEKTYFKVQVHMAPTKDEFDCKAMAYFSDV 1860
Qy 981 DLEKDVHSLIGPLLYVCHTNLNPAGROVYVOEALFETTFDEKSKYFTENMRNCRA 1040
Db 1861 DLEKDVHSLIGPLLYVCHTNLNPAGROVYVOEALFETTFDEKSKYFTENMRNCRA 1920
Qy 1041 PCNIQMEDPTEKENYRFAINCYIMDTLPGLVMAQDORIRWYLSMGSNENIHSIHFSGH 1100
Db 1921 PCNIQMEDPTEKENYRFAINCYIMDTLPGLVMAQDORIRWYLSMGSNENIHSIHFSGH 1980
Qy 1101 VFTVARKKEEYKMAIYNLYPGVFETVEMLP SKAGIWRVECLIGEHLHAQMSTLFLVYSNKC 1160
Db 1981 VFTVARKKEEYKMAIYNLYPGVFETVEMLP SKAGIWRVECLIGEHLHAQMSTLFLVYSNKC 2040
Qy 1161 QTPGLMASGHIRDPOITASGOYGQWAPKLARLHYSGSINASTKPREMIVVDLAPMI 1220
Db 2041 QTPGLMASGHIRDPOITASGOYGQWAPKLARLHYSGSINASTKPREMIVVDLAPMI 2100
Qy 1221 HGIKTOGAROKFSSLIYIQFIIMYSLDGKKNQTYRGNSGTGLMVFQVDSGSIKHNIFN 1280
Db 2101 HGIKTOGAROKFSSLIYIQFIIMYSLDGKKNQTYRGNSGTGLMVFQVDSGSIKHNIFN 2160
Qy 1281 PPIIARYIRLHPHTHYSIRSTLRMELMGCDLNSCSMPLGMEKASISDAOITASSYFTNMFA 1340
Db 2161 PPIIARYIRLHPHTHYSIRSTLRMELMGCDLNSCSMPLGMEKASISDAOITASSYFTNMFA 2220
Qy 1341 TWSPSKARLHIOGRSNARPOVNNKKEWLOVDFOKTMYVTGVTTOGVVSLTSMYVKEFL 1400
Db 2221 TWSPSKARLHIOGRSNARPOVNNKKEWLOVDFOKTMYVTGVTTOGVVSLTSMYVKEFL 2280
Qy 1401 ISSSODGHQWTLFQNGKVKYKFGQMODSFTPYVNSLDBPLLTRYLRIRIPQSVHIOIALRM 1460
Db 2281 ISSSODGHQWTLFQNGKVKYKFGQMODSFTPYVNSLDBPLLTRYLRIRIPQSVHIOIALRM 2340
Qy 1461 EVLGCEAODLY 1471
Db 2341 EVLGCEAODLY 2351

```

```

RESULT 8
US-08-366-851A-2
: Sequence 2, Application us/08366851A
: Patent No. 5681746
: GENERAL INFORMATION:
: APPLICANT: Bodner, Mordechai
: APPLICANT: De Polo, Nicolas J.
: APPLICANT: Hsu, David Chi-Tang
: APPLICANT: Chang, Steven
: TITLE OF INVENTION: Retroviral Delivery of Full Length Factor VIII

```



NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Viagene, Inc.  
STREET: 11055 Roselle Street  
CITY: San Diego  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 92121  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/366,851A  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Chambers, Daniel M.  
REGISTRATION NUMBER: 34,561  
REFERENCE/DOCKET NUMBER: 930049.438  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 452-1288  
TELEFAX: (619) 452-2616  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2351 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-366-851A-2

Query Match 94.1%; Score 7403; DB 1; Length 2351;  
Best Local Similarity 62.5%; Pred. No. 0;  
Matches 1469; Conservative 0; Mismatches 2; Indels 880; Gaps 1;  
QY 1 MOELSTFPCFLCLRFGCSATRRYYLIGAVELSMDWOSQDGLPVDARPPRYPKSFPFN 60  
DB 1 MOELSTCFCLLRFCSATRRYYLIGAVELSMDWOSQDGLPVDARPPRYPKSFPFN 60  
QY 61 TSVYVKTLFVEFTVHLFPIAKPRPMWGLGPTOAEVYDVTYTLKMMASIPVSLHAV 120  
DB 61 TSVYVKTLFVEFTVHLFPIAKPRPMWGLGPTOAEVYDVTYTLKMMASIPVSLHAV 120  
QY 121 GSVYKASEGAEYDQTSQREKEDDKVFPGSGHTYVQVLYKENGPMASDPLCTYSYLSH 180  
DB 121 GSVYKASEGAEYDQTSQREKEDDKVFPGSGHTYVQVLYKENGPMASDPLCTYSYLSH 180  
QY 121 GSVYKASEGAEYDQTSQREKEDDKVFPGSGHTYVQVLYKENGPMASDPLCTYSYLSH 180  
DB 121 GSVYKASEGAEYDQTSQREKEDDKVFPGSGHTYVQVLYKENGPMASDPLCTYSYLSH 180  
QY 181 VDLVNDLNSGLIGALLVCREGSLAKETQTLHKFILLFAVFDEGKSMHSETKNSLMQDRD 240  
DB 181 VDLVNDLNSGLIGALLVCREGSLAKETQTLHKFILLFAVFDEGKSMHSETKNSLMQDRD 240  
QY 241 AASARAWKMTHTVNGVYVRSRSLPGLIGCHRRKSYVMHYIGMGTPEVHSIFLEGTFPLVRNH 300  
DB 241 AASARAWKMTHTVNGVYVRSRSLPGLIGCHRRKSYVMHYIGMGTPEVHSIFLEGTFPLVRNH 300  
QY 301 ROASLEISPIFTFLAQTQTLMDLGOFLFCHISSHQHDMGEAVYKVDSCPEEPOLRRKKNNE 360  
DB 301 ROASLEISPIFTFLAQTQTLMDLGOFLFCHISSHQHDMGEAVYKVDSCPEEPOLRRKKNNE 360  
QY 361 EAEYDDDLTJSEMVDVVRDDNNSPSFIQIRSVAKKHKPTWVHYIAAEEDMDYADLVLA 420  
DB 361 EAEYDDDLTJSEMVDVVRDDNNSPSFIQIRSVAKKHKPTWVHYIAAEEDMDYADLVLA 420  
QY 421 PDDSKSYKSYLNNNGPQRIGRKKYKVRMAVYDETFKTRAIQIHESILLPLLYGEGDPL 480  
DB 421 PDDSKSYKSYLNNNGPQRIGRKKYKVRMAVYDETFKTRAIQIHESILLPLLYGEGDPL 480  
QY 481 LIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFPLPGEIFKYKMTVYEDGP 540  
DB 481 LIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFPLPGEIFKYKMTVYEDGP 540

QY 541 TKSDFPCLTRYISSFVNMERDLASGLIGPLLCYKESVDQRGNOIMSKRNVIYFSFDE 600  
DB 541 TKSDFPCLTRYISSFVNMERDLASGLIGPLLCYKESVDQRGNOIMSKRNVIYFSFDE 600  
QY 601 NRSWYLTENIORPLRPAGVOLEDPEFOASNTMHSINGVYPSIOLSVCLHEVAWYLLS 660  
DB 601 NRSWYLTENIORPLRPAGVOLEDPEFOASNTMHSINGVYPSIOLSVCLHEVAWYLLS 660  
QY 661 IGAQDTFLSVFSGYTFKKRMVYEDTLTFPESGETVPMSENPGLMTLIGCHNSDFNRNG 720  
DB 661 IGAQDTFLSVFSGYTFKKRMVYEDTLTFPESGETVPMSENPGLMTLIGCHNSDFNRNG 720  
QY 721 MTALLKVVSCDKNTGDIYEDSTEDISAVYLLSKNNAIEPPRSQNSRHPSTROKOPNAT 778  
DB 721 MTALLKVVSCDKNTGDIYEDSTEDISAVYLLSKNNAIEPPRSQNSRHPSTROKOPNAT 778  
QY 779 ----- 778  
DB 781 PENDLEKTPWFHAKTPMPKIQNVSSDLMLLROSPTPHGLSLDLOAKYETPSDDPS 840  
QY 779 ----- 778  
DB 841 PGALDSNNSLSEKTHRRPOLHHSQMWFTPESGDLRLNEKLGTTAATELKKIDFKVYST 900  
QY 779 ----- 778  
DB 901 SNMLSTIPSDNLAGTONTSSLGPPMPVHYDOLDPTLFGKSSPLTESGGLSLEE 960  
QY 779 ----- 778  
DB 961 NNDKSLIESGLMNSOESSGKNVSTESGRLFKGRANGPALLTKDNALFVYSISLKTN 1020  
QY 779 ----- 778  
DB 1021 KTSNNSATNRKTHIDGPSLLIENSPVMONIESDTEFFKVTPLIHDMLMDKATALRL 1080  
QY 779 ----- 778  
DB 1081 NMSNKTTSKKNMENVQAKKEGPIPPDAQNPDMSEFKMLFLPESARWIQRTHGKNSLSNG 1140  
QY 779 ----- 778  
DB 1141 QGSPKQVLSLGEKSVESQNLSEKNKYVGKGEFTADVGLKENVFPFSRNLPLTNDN 1200  
QY 779 ----- 778  
DB 1201 LHENNTNQEKKIQEIEKKEKTLIQENVLPQIHVTGTKNFMKMLFLSLTRQWVESYD 1260  
QY 779 ----- 778  
DB 1261 GAVAPYLODFRSLNDSTNRTKHTAHSKGEENLEGLGQTKQIVERYACTTRISPNT 1320  
QY 779 ----- 778  
DB 1321 SQONFVQSRKALQOFLRPLETELEKRIIVDDTSTQMSKNMKHLPTSTLTQIDYNEKE 1380  
QY 779 ----- 778  
DB 1381 KGALQPSLSDCLTRSHSIPQANRSPPLIAKVSFSPSIRPIYLRVLEFQDSSHLPAASY 1440  
QY 779 ----- 778  
DB 1441 RKKDSGVQESSHFLQAKKNNLSLAILLEMTGDQREVSLGTSATNSVYKKVENTVLP 1500  
QY 779 ----- 778  
DB 1501 KPDLPKTSKVELLPKVAHYQKDLFPTETSGSPGHLDLVBESLLOGEGAIKKNNEANRP 1560  
QY 779 ----- 778  
DB 1561 GKVPFLRVATESSAKTPSKLDPLAMDNIHGTQIQKEEMKSGEKSPEKTAKKKDTILSL 1620  
QY 779 ----- 800  
-----PVLKRRHQREITRTTIQSDOE



```

Db 1621 NACESNHAIAINEGONKPELEVWAKOGRTERLCSQNPVYLKRQREITRTTQSDDE 1680
QY 801 IDYDFTISEWKKKEPDIYDEBENSPPRSFQKTRHYTIAAVERLMDYGSSSPVHLNR 860
Db 1661 IDYDFTISEWKKKEPDIYDEBENSPPRSFQKTRHYTIAAVERLMDYGSSSPVHLNR 1740
QY 861 AOSGSVPQFKVFOEFDSFTQPLYRGELNEHLGGLPTIRAEVENINWTFPNOASR 920
Db 1741 AOSGSVPQFKVFOEFDSFTQPLYRGELNEHLGGLPTIRAEVENINWTFPNOASR 1800
QY 921 PYSFYSLSIYEEDQROGAEBRKNFVKPNETKTYFWKVQHMAPTKDEFDCKAMAYFSDV 980
Db 1801 PYSFYSLSIYEEDQROGAEBRKNFVKPNETKTYFWKVQHMAPTKDEFDCKAMAYFSDV 1860
QY 981 DLEKDVHSGLLGPLLVCTNTNLNPAHQVVOEFALEFTTJFDETKSWYFTENNERCRA 1040
Db 1861 DLEKDVHSGLLGPLLVCTNTNLNPAHQVVOEFALEFTTJFDETKSWYFTENNERCRA 1920
QY 1041 PCNTOMEPTFKENYRFAINGYIMDTLPGLVMAODRIRNYLILSMGSNENIHSHESGH 1100
Db 1921 PCNTOMEPTFKENYRFAINGYIMDTLPGLVMAODRIRNYLILSMGSNENIHSHESGH 1980
QY 1101 VETVARKKEEYKALYNLYPGVEFTEVEMLPKAGIMRVECLIGEHLHAGMSTLFLVYSNKC 1160
Db 1981 VETVARKKEEYKALYNLYPGVEFTEVEMLPKAGIMRVECLIGEHLHAGMSTLFLVYSNKC 2040
QY 1161 QTPILGMAHGIRDPQITASGOYGOMAPKLARLHSGSINANSTEPPSWIKVDLAPMI 1220
Db 2041 QTPILGMAHGIRDPQITASGOYGOMAPKLARLHSGSINANSTEPPSWIKVDLAPMI 2100
QY 1221 HGITOGAROKFESSLYISQFLIMYSLOGCKMOTYRGNSTGLMVPFGVDSGKIKHNEN 1280
Db 2101 HGITOGAROKFESSLYISQFLIMYSLOGCKMOTYRGNSTGLMVPFGVDSGKIKHNEN 2160
QY 1281 PPIIARYIRLPHYHSIRSTLRMELMGCDLNSCNPILGMEKASISDAQITASVFTNMA 1340
Db 2161 PPIIARYIRLPHYHSIRSTLRMELMGCDLNSCNPILGMEKASISDAQITASVFTNMA 2220
QY 1341 TWSPSKARLHOGSRNMAPOVNNPKEMLOVDFQKTMKVTVTTQGVKSLTSMVKEFL 1400
Db 2221 TWSPSKARLHOGSRNMAPOVNNPKEMLOVDFQKTMKVTVTTQGVKSLTSMVKEFL 2280
QY 1401 ISSSODGHOVTLFONGKVKYFGGNOGSFTPVNSLOPDLTIRLRIHPQSWHOIALRM 1460
Db 2281 ISSSODGHOVTLFONGKVKYFGGNOGSFTPVNSLOPDLTIRLRIHPQSWHOIALRM 2340
QY 1461 EVLGCBAQDLY 1471
Db 2341 EVLGCBAQDLY 2351

```

```

RESULT 9
US-07-864-004B-4
; Sequence 4, Application US/07864004B
; Patent No. 5364771
; GENERAL INFORMATION:
; APPLICANT: Lollat, John S.
; APPLICANT: Runge, Marschall S.
; TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/864,004B
; FILING DATE: 07 APRIL 1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU106
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6508
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver cDNA sequence
; US-07-864-004B-4

Query Match 93.0%; Score 7315; DB 1; Length 2332;
Best Local Similarity 62.3%; Pred. No. 0;
Matches 1452; Conservative 0; Mismatches 0; Indels 880; Gaps 1;

QY 20 ATRRYVIGAVLSMDYMSDGLPVDARFPPRPKSPFNTSVYKKTLYVEFVHLEN 79
Db 1 ATRRYVIGAVLSMDYMSDGLPVDARFPPRPKSPFNTSVYKKTLYVEFVHLEN 60
QY 80 IAKRPMPMGLGPTIQAEVYDVTYITLKMAHSPVSLHAGVSYWKASEGAEVDDOTSQ 139
Db 61 IAKRPMPMGLGPTIQAEVYDVTYITLKMAHSPVSLHAGVSYWKASEGAEVDDOTSQ 120
QY 140 REKEDDKVFPGSGHTYVQVLKENGPMASDPLCTYSLSHVLDLVKDNSLIGALLVCR 199
Db 121 REKEDDKVFPGSGHTYVQVLKENGPMASDPLCTYSLSHVLDLVKDNSLIGALLVCR 180
QY 200 EGSIAKETQTLHFIILFAVFDGKSMHSTKNSLMQDRDAASARAPKMTYNGYNR 259
Db 181 EGSIAKETQTLHFIILFAVFDGKSMHSTKNSLMQDRDAASARAPKMTYNGYNR 240
QY 260 SLPLGIGCHRKSYWVHYGMCTPEVHSIFLEGHFVLRNHRQASLEISPTTFLTAOTLL 319
Db 241 SLPLGIGCHRKSYWVHYGMCTPEVHSIFLEGHFVLRNHRQASLEISPTTFLTAOTLL 300
QY 320 MDLGOFLFCHISSHQHDMGMEAVYKVDSCPEPOLRMKNNEADYDDDLTDEMDVYRF 379
Db 301 MDLGOFLFCHISSHQHDMGMEAVYKVDSCPEPOLRMKNNEADYDDDLTDEMDVYRF 360
QY 380 DDDNSPFIQIRSAKHKPKWVHYIAAEEEDMDYAPLYLAPDDRYSKSOYLNNQPORIG 439
Db 361 DDDNSPFIQIRSAKHKPKWVHYIAAEEEDMDYAPLYLAPDDRYSKSOYLNNQPORIG 420
QY 440 RKYKVFEMATYDTEFTKREAIQIHESGLIGPLLGEVGDITLLIFKNQASPPYNIYPHGI 499
Db 421 RKYKVFEMATYDTEFTKREAIQIHESGLIGPLLGEVGDITLLIFKNQASPPYNIYPHGI 480
QY 500 TDVRPLYSRLPKGVKHLKDEPILPGEIFKYKWTVTVEDGPTKSDPCLTRYSSFYVME 559
Db 481 TDVRPLYSRLPKGVKHLKDEPILPGEIFKYKWTVTVEDGPTKSDPCLTRYSSFYVME 540
QY 560 RDLASGLIGPLLCYKESVQGRNOIMSDKRNVLFSVPDENRSWYLTENTQRLPLPAG 619
Db 541 RDLASGLIGPLLCYKESVQGRNOIMSDKRNVLFSVPDENRSWYLTENTQRLPLPAG 600
QY 620 VQLEDPEFOASNIHNSINGVYFDSIQSLVCLHEVAAYWTLSIGAQDTFLSVFFSGYTFKH 679

```

Db 601 VOLEDPFQASNIHMSINGYVFDLSQLSVCLHEVAWYILSLGAQOTDPLSVFFSGYTFKH 660  
QY 680 KMVYEDTLTLPFSGEIVEMSMENGLWILGCHNSDFRNRGMTALLKVSCKDKNTGDYXE 739  
Db 661 KMVEDITLTLFPFSGEIVEMSMENGLWILGCHNSDFRNRGMTALLKVSCKDKNTGDYXE 720  
QY 740 DSYEDISAYLLSKNNAIEPRSFQNSRHPSTROKOPNAT 778  
Db 721 DSYEDISAYLLSKNNAIEPRSFQNSRHPSTROKOPNAT 780  
QY 779 778  
Db 781 KIQNVSSDLMMLRQSPFPHGLSLSDLOEAKYETFPSPDPSGALDSNNSLSEMTHERPO 840  
QY 779 778  
Db 841 LHHSGDMVFTPEBSGLQLRLNEKLTGTAATELKLDKFKVSTSNLSTIPSDNLAAGTON 900  
QY 779 778  
Db 901 TSLGPPSPVHYDSQDLDTLTFGKKSPLTESGGLSLSENNDSKLESGLMNSQESSM 960  
QY 779 778  
Db 961 GKNVSTESGRLFKGRAHPALLTKDNALFKVSLSLKTNTSNNATNRKTHIDGPSL 1020  
QY 779 778  
Db 1021 LIENSPSWQNLIESDTEFKKYTPLIHDRMLMDKNATALRLHMSNKTSSKNMEVQOK 1080  
QY 779 778  
Db 1081 KEPIPPDAQNPDMSEFKMLFLPESARMIQRTHGKNSLNSGQSPKQLVSLGPEKSVEG 1140  
QY 779 778  
Db 1141 QNPLSEKKNVVVGKGTQVGLKEMPEPSSRNLFPLNDLHNNTHNOEKKIOEIEIK 1200  
QY 779 778  
Db 1201 KETLIGENVVLPQIHVYGTGKNFMKNLPLSTRQNVESGEGAYAPVLODFSLNDSTR 1260  
QY 779 778  
Db 1261 TKKHTAHSKKGEEENLEGLNOTQIOVEKACTTRISPNTSQONEVTORSKRALKQFRL 1320  
QY 779 778  
Db 1321 PLEETELEKRIIVDDTSTQWSKNMKHLTPSYLTQIDYNEKEGALTQSPLSDCLTRSHSI 1380  
QY 779 778  
Db 1381 PQANRSPPLIAKVSPSIRPIYTLRVLEFQDNSSHLPASRYKKDSGVQESSHFLQGA 1440  
QY 779 778  
Db 1441 NNLSLAILTEMTGDQREVSLGTSATNSVTYKVENFVLPKPDLPKTSKVYELLPKVHI 1500  
QY 779 778  
Db 1501 YOKDLFTETSGNSPGHLDLVEGSLLOGTEGAIKWNEANRPGKVPFLRVATESAKTPSK 1560  
QY 779 778  
Db 1561 LLDPLANDNHGTQIPKEEMKSQEKSPETAKKKDTILSLACESNNAIAINEGONKP 1620  
QY 779 819  
Db 1621 EIEVTMAQGRTERLCSQNPVYLRHOREITRTLOSQDEIDYDPTISEVKKEDFDY 1680  
QY 820 DEDENOSRSPQOKTRHYFIAAVERLMDYGMSSSPHYLBNRQSGSVPOFKVYVFOEFTD 879  
Db 1681 DEDENOSRSPQOKTRHYFIAAVERLMDYGMSSSPHYLBNRQSGSVPOFKVYVFOEFTD 1740

QY 880 GSETOPLYRGELNEHGLGPIYRAEVEDNIMVTERNOASRPYSLSLISEEDORGA 939  
Db 1741 GSFTOPLYRGELNEHGLGPIYRAEVEDNIMVTERNOASRPYSLSLISEEDORGA 1800  
QY 940 EPRKNEVAKNEFKTYFWKQHHMAPTKDEFDCKAANYISDVLEKDVHSGILGPLYCHT 999  
Db 1801 EPRKNEVAKNEFKTYFWKQHHMAPTKDEFDCKAANYISDVLEKDVHSGILGPLYCHT 1860  
QY 1000 NTLNPAHGRQVYVQEFALFEFTJFDETKSWYFTENNERNCRAPCNQOMDPTFKENYRPHA 1059  
Db 1861 NTLNPAHGRQVYVQEFALFEFTJFDETKSWYFTENNERNCRAPCNQOMDPTFKENYRPHA 1920  
QY 1060 INGYIMDTLPGLVMAODORIRRYLLSMGSNENHSHIRSGHVFYVRKKEEYKMALVNLXP 1119  
Db 1921 INGYIMDTLPGLVMAODORIRRYLLSMGSNENHSHIRSGHVFYVRKKEEYKMALVNLXP 1980  
QY 1120 GVEFTEVEMLPKAGIWRBCLIGEHLHAGMSTFLVYSNKCOTPLGMASGHIRFOITAS 1179  
Db 1981 GVEFTEVEMLPKAGIWRBCLIGEHLHAGMSTFLVYSNKCOTPLGMASGHIRFOITAS 2040  
QY 1180 GOYGOAPKLARLHYSGSINAMSTKEPFSWIKVDLLAPMITHGIRTOGAROKFSSLYTSQ 1239  
Db 2041 GOYGOAPKLARLHYSGSINAMSTKEPFSWIKVDLLAPMITHGIRTOGAROKFSSLYTSQ 2100  
QY 1240 FIIMYSLDGKKMOTYRGNSTGTLMWFFGNVDSGJIKHNIENPPIARYIRLPHYSIRS 1299  
Db 2101 FIIMYSLDGKKMOTYRGNSTGTLMWFFGNVDSGJIKHNIENPPIARYIRLPHYSIRS 2160  
QY 1300 TIRMEIMGCDLNSCNPUGMESKAISDAQITASSFTWKFATWSPSKRHLQGRSNAMR 1359  
Db 2161 TIRMEIMGCDLNSCNPUGMESKAISDAQITASSFTWKFATWSPSKRHLQGRSNAMR 2220  
QY 1360 POVNNPKEMLOYDFOKTKMYGVTTQGVKSLTSMYKVEFLISSQDSQDHQWTLFFQNGKV 1419  
Db 2221 POVNNPKEMLOYDFOKTKMYGVTTQGVKSLTSMYKVEFLISSQDSQDHQWTLFFQNGKV 2280  
QY 1420 KVEQGNDSFTPVNSLDPPLLTRYLRIHPQSWHQIALRMEVLCGEAODLY 1471  
Db 2281 KVEQGNDSFTPVNSLDPPLLTRYLRIHPQSWHQIALRMEVLCGEAODLY 2332  
  
RESULT 10  
US-08-251-937A-4  
; Sequence 4, Application US/08251937A  
; Patent No. 5583209  
; GENERAL INFORMATION:  
; APPLICANT: Lollar, John S.  
; APPLICANT: Runge, Marschall S.  
; TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Kilpatrick & Cody  
; STREET: 1100 Peachtree Street  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: US  
; ZIP: 30309  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/251,937A  
; FILING DATE: 31-MAY-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/864,004  
; FILING DATE: 07-APR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pratt, John S.  
; REGISTRATION NUMBER: 29,476

REFERENCE/DOCKET NUMBER: EMU106DIV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-815-6367  
TELEFAX: 404-815-6555  
INFORMATION FOR SEO ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2332 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Homo sapien  
TISSUE TYPE: Liver cDNA sequence  
US-08-251-937A-4

Query Match 93.0%; Score 7315; DB 1; Length 2332;

Best Local Similarity 62.3%; Pred. No. 0;  
Matches 1452; Conservative 0; Mismatches 0; Indels 880; Gaps 1;

QY 20 ATRRYLGAVELSMDYMOSDLGELPVDARPPRPVKSPFNTSVYKKTLFVEFTVHLFN 79  
DB 1 ATRRYLGAVELSMDYMOSDLGELPVDARPPRPVKSPFNTSVYKKTLFVEFTVHLFN 60  
QY 80 IAKRPPMGLGPTIOAEVYDVVVITLKNNASHPVSLHAGVSYWKASEGAEYDDQTSQ 139  
DB 61 IAKRPPMGLGPTIOAEVYDVVVITLKNNASHPVSLHAGVSYWKASEGAEYDDQTSQ 120  
QY 140 REKEDKAFPEGSHTYVQVLAKEGPMASDPLCTLYSLSHVDLVKDLNSGLIGALLVCR 199  
DB 121 REKEDKAFPEGSHTYVQVLAKEGPMASDPLCTLYSLSHVDLVKDLNSGLIGALLVCR 180  
QY 200 EGSIAKEKTOFLHKFILLFAVDEGKSWHSETKNSLMODRPAASARAMPKMTVNGVYNR 259  
DB 181 EGSIAKEKTOFLHKFILLFAVDEGKSWHSETKNSLMODRPAASARAMPKMTVNGVYNR 240  
QY 260 SLPLGICGRKSVYWHVIGMGTTPVHSLFLEGHTFLVNRHQALESIPITFLTAQTL 319  
DB 241 SLPLGICGRKSVYWHVIGMGTTPVHSLFLEGHTFLVNRHQALESIPITFLTAQTL 300  
QY 320 MDLQGFLLFCHSHOHGMEAYYVDCSPREPQLRMKNNEAEYDDDLTDSMDVYRF 379  
DB 301 MDLQGFLLFCHSHOHGMEAYYVDCSPREPQLRMKNNEAEYDDDLTDSMDVYRF 360  
QY 380 DDDNSPFIQIRSAVAKKHPKTMVHYIAAEEDMDAPLVLPDDRYSKSOYLNNGPORIG 439  
DB 361 DDDNSPFIQIRSAVAKKHPKTMVHYIAAEEDMDAPLVLPDDRYSKSOYLNNGPORIG 420  
QY 440 RKKYKVRMAVYDDETFKTREALIQHESGILGPLLYGEVDTLLIFKNQASRPYNIYPHGI 499  
DB 421 RKKYKVRMAVYDDETFKTREALIQHESGILGPLLYGEVDTLLIFKNQASRPYNIYPHGI 480  
QY 500 TDVAPLVRRLPKVGYKHLKDPILLPELFFKYKWTYVDEGPTKSDPRLTRYYSFVME 559  
DB 481 TDVAPLVRRLPKVGYKHLKDPILLPELFFKYKWTYVDEGPTKSDPRLTRYYSFVME 540  
QY 560 RDLASGLGPLLICYSKESVDORGNQIMSDKRNVLIFSVEENRSYULENTORFLPNPAG 619  
DB 541 RDLASGLGPLLICYSKESVDORGNQIMSDKRNVLIFSVEENRSYULENTORFLPNPAG 600  
QY 620 VQLEDPFQASNIHMSINGVYFDSLQTSVCLHEVAWYILSIGAQDTLVSFFSGYTERKH 679  
DB 601 VQLEDPFQASNIHMSINGVYFDSLQTSVCLHEVAWYILSIGAQDTLVSFFSGYTERKH 660  
QY 680 KMYEDDTLLPPEGGEYFEMENPGMLIILCHNSDFRNKRGMTALIKSSCDKNTGDIYE 739  
DB 661 KMYEDDTLLPPEGGEYFEMENPGMLIILCHNSDFRNKRGMTALIKSSCDKNTGDIYE 720  
QY 740 DSYEDISAYLLSKNNALIPRSFSQNSRHPSTROKOFNAT----- 778

DB 721 DSYEDISAYLLSKNNALIPRSFSQNSRHPSTROKOFNATTTIPENDIEKTDPAHRTMP 780  
QY 779 ----- 778  
DB 781 KIONVSSDILLMLRQSPHPHGLSLSLQEAKEYTFSDPSGAIDSNNSISEMTHFRPQ 840  
QY 779 ----- 778  
DB 841 LHHSGDMVFTPEGLQLRLNEKLGTTAATELKKLDKVSSTSNLITSTPSDNLAAGTDN 900  
QY 779 ----- 778  
DB 901 TSSLGPPMPVHYDSQDLDTTLFGKSSPLTSGGPTLSSEENDSKLLESGILMNSQESSW 960  
QY 779 ----- 778  
DB 961 GKNVSTESGLRFGKRAHGALLTKDNALFKVISLTKTNKTSNNSATNKRTHIDGPSL 1020  
QY 779 ----- 778  
DB 1021 LIENSPYWMNILESDEFEKKVPTLIHDMMLDKNATRLHMSNKTSSKNMEMVOOK 1080  
QY 779 ----- 778  
DB 1081 KEGTLPDAQNDMSFFKMLFLPESARWIOPTHGKNSLNSGGGSPKQOLVSLGPEKSYEG 1140  
QY 779 ----- 778  
DB 1141 QNFLSEKKNVYVKGGEFTKDVGLKEMVFPSSRNFLTLNLDNLHENTHNOEKKIOEBIEK 1200  
QY 779 ----- 778  
DB 1201 KETLLQENVVLPQIHTVTGTRKFNKNLFLSTRONVGSYEGAVAPVLODFRSLDSTNR 1260  
QY 779 ----- 778  
DB 1261 TKKTHAHSKSGEENLGLNOTQOYIEKACTTRISPNSQGFVYQSRKALKQFRL 1320  
QY 779 ----- 778  
DB 1321 PLEETELEKRIIVDDTSTQMSKNMKHLTPSTLQIDYNEKEGAIQTGSLDCLTRSHSI 1380  
QY 779 ----- 778  
DB 1381 PQANRSLPLIAKVSSPISIRPIYLTRVLPQDNSSHLPASTRKDSGVQESSHFLQAKK 1440  
QY 779 ----- 778  
DB 1441 NNLSLAILTLEMTGQREVSGLSGTSATNSVYKKYENTVLPKPDLPKTSKVELLPKYHI 1500  
QY 779 ----- 778  
DB 1501 YOKDLFPTETSGNSPGHLDIVEGSLLOSTEGAIKWNEANRPGKVPFLRVATESSAKTPSK 1560  
QY 779 ----- 778  
DB 1561 LLDPLAMDNHYGTQIPKEEMKSQEKSPKTAFFKKOTLLSNACESNHAIAINEGQNK 1620  
QY 779 ----- 778  
DB 1621 ELEVYMAOGTRELCSQNPVYLRKHQREITRTTIQSQOEIYDDDTISVEKKKEDFPIT 1680  
QY 820 DEDENQSPRSFOKTI RHFIYAVERLMDYGMSSSPHYLRNBAQSGSVQFKVYFOEFTD 879  
DB 1681 DEDENQSPRSFOKTI RHFIYAVERLMDYGMSSSPHYLRNBAQSGSVQFKVYFOEFTD 1740  
QY 880 GSFQPLVRCGLNEHLGLGPIYIRAEVEDNIMVPRNQAASRPYSGLISYEEOORGA 939  
DB 1741 GSFQPLVRCGLNEHLGLGPIYIRAEVEDNIMVPRNQAASRPYSGLISYEEOORGA 1800  
QY 940 EPRKNFVKNETKTYFWKVQHMAPTKDEFDCKAMAYFSDVLEKDVSGILGPLVCHT 999

```

Db 1801 EPRKAEVKEKNETKTYFWKVOHHMAPTKDEBDCKAMAYFSDVDLEKDVHSLGILPVLVCHT 1860
QY 1000 NTLNPAHGRQVTVQEBALFFTFIDETKSWYFTENMERNCRAPNIOMEDTEFKENYRFAH 1059
Db 1861 NTLNPAHGRQVTVQEBALFFTFIDETKSWYFTENMERNCRAPNIOMEDTEFKENYRFAH 1920
QY 1060 INGYIMDTLPGLVMAODQIRIWMYLLSMGSNENIHSHFSGHVFYTKKEEYKMALYNLYP 1119
Db 1921 INGYIMDTLPGLVMAODQIRIWMYLLSMGSNENIHSHFSGHVFYTKKEEYKMALYNLYP 1980
QY 1120 GVEFTEVEMPSKAGIMRWVCLIGEHLLHAGMSTFLVYSNKCQPLCMASGHINDPQITAS 1179
Db 1981 GVEFTEVEMPSKAGIMRWVCLIGEHLLHAGMSTFLVYSNKCQPLCMASGHINDPQITAS 2040
QY 1180 GYGGMAPKLARLHYSGSINAMSTKEPFSMIKVDLAPMIIHIGIKTOGAROKFSSYIISO 1239
Db 2041 GYGGMAPKLARLHYSGSINAMSTKEPFSMIKVDLAPMIIHIGIKTOGAROKFSSYIISO 2100
QY 1240 FIIMYSLDGKKWOTYRGNSTGTLLMVEFGVNDSSGIKHNIFNPPIIARYIRLAPTHYSINS 1299
Db 2101 FIIMYSLDGKKWOTYRGNSTGTLLMVEFGVNDSSGIKHNIFNPPIIARYIRLAPTHYSINS 2160
QY 1300 TLRMLMGCDLNSGSMPLGMEKSAISDAQITASSYFTNMFATWSPSKARLHLOGRGNAMR 1359
Db 2161 TLRMLMGCDLNSGSMPLGMEKSAISDAQITASSYFTNMFATWSPSKARLHLOGRGNAMR 2220
QY 1360 POVNNPKEMLDQVDFQTKMYVTGTTQGVKSLTSMYKKEFLISSODGQHWTLFPONGKY 1419
Db 2221 POVNNPKEMLDQVDFQTKMYVTGTTQGVKSLTSMYKKEFLISSODGQHWTLFPONGKY 2280
QY 1420 KVFQGNOSFTFVNSLDPPLLTRILRIHPQSVHQAIALRMEVLGCEADLY 1471
Db 2281 KVFQGNOSFTFVNSLDPPLLTRILRIHPQSVHQAIALRMEVLGCEADLY 2332

RESULT 11
US-08-212-133A-2
: Sequence 2, Application US/08212133A
: Patent No. 5663060
: GENERAL INFORMATION:
: APPLICANT: Lollar, John S.
: APPLICANT: Runge, Marshall S.
: TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Kilpatrick & Cody
: STREET: 100 Peachtree Street
: City: Atlanta
: STATE: Georgia
: COUNTRY: US
: ZIP: 30303
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/212.133A
: FILING DATE: March 11, 1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/864,004
: FILING DATE: 07-APR-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Pabst, Patrea L.
: REGISTRATION NUMBER: 31,284
: REFERENCE/DOCKET NUMBER: EMU/76677
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 404-572-6555
: TELEFAX: 404-572-6555
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2332 amino acids

```

```

: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHEICAL: YES
: AMTI-SENSE: NO
: FRAGMENT TYPE: N-terminal
: ORIGINAL SOURCE:
: ORGANISM: Homo sapien
: TISSUE TYPE: Liver cDNA sequence
: US-08-212-133A-2

```

```

Query Match          93.0%; Score 7315; DB 1; Length 2332;
Best Local Similarity 62.3%; Pred. No. 0;
Matches 1452; Conservative 0; Mismatches 0; Indels 880; Gaps 1;

```

```

QY 20 ATRRRYLGAVELSDWYMGSDGLBLYDARFPPRVKSPFPNTSVYKKTLVEFTVHLFN 79
Db 1 ATRRRYLGAVELSDWYMGSDGLBLYDARFPPRVKSPFPNTSVYKKTLVEFTVHLFN 60
QY 80 IAKRPPMMGLGPTIOAEVYDTVYTTLKNMASHVSLHANGSVYWKASGEAYDQTSO 139
Db 61 IAKRPPMMGLGPTIOAEVYDTVYTTLKNMASHVSLHANGSVYWKASGEAYDQTSO 120
QY 140 REKEDKVEPPGSHYVWOLKENGPMASDPLCTYSYLSHVLDVYKDLNSGLIGALVCR 199
Db 121 REKEDKVEPPGSHYVWOLKENGPMASDPLCTYSYLSHVLDVYKDLNSGLIGALVCR 180
QY 200 EGSIAKEKTQTLAKFTLLFAVDEGKSWHSFTKNSLMODROAASRAMPKMTVNGVYNR 259
Db 181 EGSIAKEKTQTLAKFTLLFAVDEGKSWHSFTKNSLMODROAASRAMPKMTVNGVYNR 240
QY 260 SLPGILGCHRRKSYVWHVIGMGTTPVHSIFLEGHTFLVNRHROASLESPTFLTAQTL 319
Db 241 SLPGILGCHRRKSYVWHVIGMGTTPVHSIFLEGHTFLVNRHROASLESPTFLTAQTL 300
QY 320 MDLGQFLFCCHISSHQDGMAYVYKVDSCPEEPOLRMKNNEBAEDYDDDLTDSMDVYRF 379
Db 301 MDLGQFLFCCHISSHQDGMAYVYKVDSCPEEPOLRMKNNEBAEDYDDDLTDSMDVYRF 360
QY 380 DDDNSPSFIQISVAKKHRTVHTYIAAEEEDWDYAPLYLADDDNSYSQYILNKPQIRG 439
Db 361 DDDNSPSFIQISVAKKHRTVHTYIAAEEEDWDYAPLYLADDDNSYSQYILNKPQIRG 420
QY 440 RRYKKVRFMAVYDEPFTKEATOHESGILGPLYGEVGTLLITRKNQASPRYNTYPRGI 499
Db 421 RRYKKVRFMAVYDEPFTKEATOHESGILGPLYGEVGTLLITRKNQASPRYNTYPRGI 480
QY 500 TDVRLYSRRLPRGVKHLDPILPGLIFRYKMTVYVEDGPTKSDPRLTRYSSFVME 559
Db 481 TDVRLYSRRLPRGVKHLDPILPGLIFRYKMTVYVEDGPTKSDPRLTRYSSFVME 540
QY 560 ROLASGLIPPLICYKESVDQGNQIMSDKRVYILFVSFDEKRSYLTENIQRFIPNAG 619
Db 541 ROLASGLIPPLICYKESVDQGNQIMSDKRVYILFVSFDEKRSYLTENIQRFIPNAG 600
QY 620 VOLEDEFOASNIMHSINGVYDSIQLSVCLHEVAWYIISGAOTDLSVFFSGYTKH 679
Db 601 VOLEDEFOASNIMHSINGVYDSIQLSVCLHEVAWYIISGAOTDLSVFFSGYTKH 660
QY 680 KMYVEDTLTLFPFSGEYFVSMENPGLWILGCHNSDFRNKGTALLKVSCKDKNTGYYE 739
Db 661 KMYVEDTLTLFPFSGEYFVSMENPGLWILGCHNSDFRNKGTALLKVSCKDKNTGYYE 720
QY 740 DSYEDISATLISKNNAIERSFSQNSRHPSTRQOKPNAI----- 778
Db 721 DSYEDISATLISKNNAIERSFSQNSRHPSTRQOKPNAITTPENDIEKTDPMFAHRTPM 780
QY 779 ----- 778
Db 781 KIQNVSSSDLLMLLRQSPPHGLSLSDIQEAKYETTFSDPSGAIIDNSNLSSEWTHRRPQ 840

```

```

QY 779 ----- 778
Db 841 LHHSGDWFTPESSGLQLRLNEKLTGTAATELKLDKFKVSVTSNNLISITPSDNLACTDN 900
QY 779 ----- 778
Db 901 TSSLOPSPMVAHYDSQDLDTLTFGKKSSPLTESGGLSLSEBNDKSKLESGLMNSQESSW 960
QY 779 ----- 778
Db 961 GKNVSTESGRLFKGRAHGPALLTKDNALFKVSTISLKTNTKSNNKATNKKTHIDGDSL 1020
QY 779 ----- 778
Db 1021 LIENSPSWQNLIESDPEFKVTPLIHDRMLMDKNATALRLNHSNKTTSSKNEMVQOK 1080
QY 779 ----- 778
Db 1081 KEGPIPPDAQNDMSFFKMLFLPESARIQRTHGKNSLNSGQSPKQVSLGPEKSVEG 1140
QY 779 ----- 778
Db 1141 QNPLSEKNKVVVGGEFTKDVGLKEMVPPSSRNLELTNLDNLHENNTHNOEKKIOEIEK 1200
QY 779 ----- 778
Db 1201 KETLIGENVVLPQIHVTYGTNKNMKNLFLSTRONVSGSYEGATAPVLQDFRSLNDSTNR 1260
QY 779 ----- 778
Db 1261 TKKHTAHSKKEEENLEGLQNTQKQIVEKYACTTRISPTNSQONFVYQSRKALKQFRL 1320
QY 779 ----- 778
Db 1321 PLEETELEKRIYDDTSTQMSKMKHLTPSTLQIDYNEKEKAITQSPSLDCLTRSHSI 1380
QY 779 ----- 778
Db 1381 PQANRSLPLAKVSSPFSIRIYLTRLVLFQDNSSHLPASVYRKKGSGVQESHHFQCAKK 1440
QY 779 ----- 778
Db 1441 NNLSLAILTEMTGDQREVGSLGTSATNSVYTKKVENTVLPKPDLPKTSKVELLPKVIH 1500
QY 779 ----- 778
Db 1501 YOKDLPTETSNSGPHLDIVEGSLLOGTEGAIKMNEANRPKVPFLKVAATESAKTPSK 1560
QY 779 ----- 778
Db 1561 LDDPLAMDNHGTQIPKEEMKSQEKSPKTAFFKKDTILSNACESNHAIAINEGONKP 1620
QY 779 ----- 819
Db 1621 EIEVTWAKQRTERLCSQNPVPLKRHOEIRITLQSDQEIYDDTISVMKKEDFDIY 1680
QY 820 DEDNOSPRSKOKTRHFIYAVERLDYGSSPHVLRNAGSGSVQFKKVPQOETD 879
Db 1681 DEDNOSPRSKOKTRHFIYAVERLDYGSSPHVLRNAGSGSVQFKKVPQOETD 1740
QY 880 GSFQPLRYGELNHLGLGPIYIRAEVEDNIMVTFRNQASRPYSFYSSLIYEEDORGA 939
Db 1741 GSFQPLRYGELNHLGLGPIYIRAEVEDNIMVTFRNQASRPYSFYSSLIYEEDORGA 1800
QY 940 EPRKNFYKPNFTKTYFKVQHMAPTYKDEPCKAMAFSDVDLEKDVHSGLIGLLVCHT 999
Db 1801 EPRKNFYKPNFTKTYFKVQHMAPTYKDEPCKAMAFSDVDLEKDVHSGLIGLLVCHT 1860
QY 1000 NTLNPAHGRQVTOEFALEFIIPEETKSWYFTEEMNERCRAPCNOMDEPFKKNYRPHA 1059
Db 1861 NTLNPAHGRQVTOEFALEFIIPEETKSWYFTEEMNERCRAPCNOMDEPFKKNYRPHA 1920
QY 1060 INGYIMDTLPGLVMAODORIRMYILMSGNSNENIHSIHSGHVFTRKKEEKMALYNLYP 1119

```

```

Db 1921 INGYIMDTLPGLVMAODORIRMYILMSGNSNENIHSIHSGHVFTRKKEEKMALYNLYP 1980
QY 1120 GVFETVEMLPESKAGIWRRECELIGELHAGMSTFLYVSNKOTPLGMAAGHIRPQITAS 1179
Db 1981 GVFETVEMLPESKAGIWRRECELIGELHAGMSTFLYVSNKOTPLGMAAGHIRPQITAS 2040
QY 1180 GYGOMAPKRLARLHYSGSINMSTKEPFSWTKVDLAPMIITHGIKTGAROKFSSLYTSO 1239
Db 2041 GYGOMAPKRLARLHYSGSINMSTKEPFSWTKVDLAPMIITHGIKTGAROKFSSLYTSO 2100
QY 1240 FIIMYSLDGKKWQTYRGNSTGLTWFFGNVDSGKIHAFNPPIIARIIRLPHYSIRS 1299
Db 2101 FIIMYSLDGKKWQTYRGNSTGLTWFFGNVDSGKIHAFNPPIIARIIRLPHYSIRS 2160
QY 1300 TLRMELMGCDLNSCSMPJGMSKASIDAOITASSYFTNMFATWSPSKARLHOGRSNMR 1359
Db 2161 TLRMELMGCDLNSCSMPJGMSKASIDAOITASSYFTNMFATWSPSKARLHOGRSNMR 2220
QY 1360 POWNPKEMLOVDFOKTKMYGVTTQGVKSILTSYVKEEPLISSODGHQMTLFFQNGV 1419
Db 2221 POWNPKEMLOVDFOKTKMYGVTTQGVKSILTSYVKEEPLISSODGHQMTLFFQNGV 2280
QY 1420 KVFQGNDSFTPPVNSLDPPLTRYLRIHPOSWVHQIALRMEVLGCEAODLY 1471
Db 2281 KVFQGNDSFTPPVNSLDPPLTRYLRIHPOSWVHQIALRMEVLGCEAODLY 2332

```

```

RESULT 12
US-08-474-503-2
: Sequence 2, Application US/08474503
: Patent No. 5744446
:
: GENERAL INFORMATION:
: APPLICANT: Emory University
: TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Kilpatrick & Cody
: STREET: 1100 Peachtree Street, Suite 2800
: CITY: Atlanta
: STATE: Georgia
: COUNTRY: US
: ZIP: 30309
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/474,503
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Pratt, John S.
: REGISTRATION NUMBER: 29,476
: REFERENCE/DOCKET NUMBER: EMU106CIP(3)
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 404-815-6500
: TELEFAX: 404-815-6555
:
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2332 amino acids
: type: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHEICAL: YES
: ANTI-SENSE: NO
: FRAGMENT TYPE: N-terminal
: ORIGINAL SOURCE:
: ORGANISM: Homo sapien
: TISSUE TYPE: Liver cDNA sequence
:
: US-08-474-503-2

```

Query Match 93.0%: Score 7315; DB 1: Length 2332;  
 Best Local Similarity 62.3%: Pred. No. 0;  
 Matches 1452: Conservative 0; Mismatches 0; Indels 880; Gaps 1;

QY 20 ATRRYLGAVELSMDYMSDGLGELPVDARFPPRPVKSPFPNTSVYKKTLLFVEFTVHLEN 79  
 |||||  
 DB 1 ATRRYLGAVELSMDYMSDGLGELPVDARFPPRPVKSPFPNTSVYKKTLLFVEFTVHLEN 60

QY 80 IAKRRPMMGLGPTIOAEVDTVYITLKMAHSPVSLHAGVSVKASGEAGYDQTSQ 139  
 |||||  
 DB 61 IAKRRPMMGLGPTIOAEVDTVYITLKMAHSPVSLHAGVSVKASGEAGYDQTSQ 120

QY 140 REKEDDKVPGSGSHYVWQVLKENGPMASDPLCTVSYLSHVDLVKDLNSGIGALLVCR 199  
 |||||  
 DB 121 REKEDDKVPGSGSHYVWQVLKENGPMASDPLCTVSYLSHVDLVKDLNSGIGALLVCR 180

QY 200 EGSIAKEKQTLHKFILLFAVFDEGKSMHSETKNSLMDORDAASARAMPKMTVNGYVR 259  
 |||||  
 DB 181 EGSIAKEKQTLHKFILLFAVFDEGKSMHSETKNSLMDORDAASARAMPKMTVNGYVR 240

QY 260 SLPLGICGRKSVYWHVIGMGTPEVHSLFLEGHTEFLVNNHQAQSLSPITFLTAQTLL 319  
 |||||  
 DB 241 SLPLGICGRKSVYWHVIGMGTPEVHSLFLEGHTEFLVNNHQAQSLSPITFLTAQTLL 300

QY 320 MDLGOFLLFCHISSHQHDMGEAYVVDSCPEEPOLRMKNBEAEADYDDDLTSEMDVYRF 379  
 |||||  
 DB 301 MDLGOFLLFCHISSHQHDMGEAYVVDSCPEEPOLRMKNBEAEADYDDDLTSEMDVYRF 360

QY 380 DDDNSPSEIQRISVAKKHPTWVHYIAEEDMDYAPLVLPADDRSYKQYLNNQPORG 439  
 |||||  
 DB 361 DDDNSPSEIQRISVAKKHPTWVHYIAEEDMDYAPLVLPADDRSYKQYLNNQPORG 420

QY 440 RKYKVRMATDDETFKREAIQHESGILGPLLYGEGDTLLIIRKQASRPYNTYPHCI 499  
 |||||  
 DB 421 RKYKVRMATDDETFKREAIQHESGILGPLLYGEGDTLLIIRKQASRPYNTYPHCI 480

QY 500 TDVRLPLSRLLPKGVKHLKDFPILPGEIFKYKMTVTEDEGPTKSDPRLCTRYSSFVME 559  
 |||||  
 DB 481 TDVRLPLSRLLPKGVKHLKDFPILPGEIFKYKMTVTEDEGPTKSDPRLCTRYSSFVME 540

QY 560 RDLASGLIGPLLCYKESVDORGNQIMSDKRVNLLFSVDENRSMYLTENIOFLPNPAG 619  
 |||||  
 DB 541 RDLASGLIGPLLCYKESVDORGNQIMSDKRVNLLFSVDENRSMYLTENIOFLPNPAG 600

QY 620 VOLDEPBOANIMHSINGVYFDSLOLSVCLHEVAYWYILSIGAOTDPLSVFSGTFFKH 679  
 |||||  
 DB 601 VOLDEPBOANIMHSINGVYFDSLOLSVCLHEVAYWYILSIGAOTDPLSVFSGTFFKH 660

QY 680 KMYEDTLTLPFSGETVFMSEMENPGLWILGCHNSDFRNRGWTALLKYSKCDKNTGDYXE 739  
 |||||  
 DB 661 KMYEDTLTLPFSGETVFMSEMENPGLWILGCHNSDFRNRGWTALLKYSKCDKNTGDYXE 720

QY 740 DSYEDISAYLLSKNNALIPRSFSQNSRHPSTRQKQFNAT - - - - - 778  
 |||||  
 DB 721 DSYEDISAYLLSKNNALIPRSFSQNSRHPSTRQKQFNAT - - - - - 780

QY 779 - - - - - 778  
 |||||  
 DB 781 KIONVSSDMLMLRQSPTPHGLSLSLQEAKEYTFSDPSGAIDSNNSLSEMTHFRRQ 840

QY 779 - - - - - 778  
 |||||  
 DB 841 LHHSGDMVFTPEBGLQLRLNEKLGTTAATLKKLDFKVSSTSNLITSTIPDNLAAGTGN 900

QY 779 - - - - - 778  
 |||||  
 DB 901 TSSLGPPSMPIVHYDSQDLDTLFGKSSPLTESGPLSLSEENNDKSLSEGLMNSQESSM 960

QY 779 - - - - - 778  
 |||||  
 DB 961 GKNVSTESGRLFKGRAHPALLTKNALFKVSIISLKTNKTNNSATNRKTHIDGPSL 1020

QY 779 - - - - - 778  
 |||||  
 DB 1021 LIENSPSWQNLLESQTEKKKVTPLIHDMMLDKNATALRLNHMSKNTTSSKNEMWQOK 1080

QY 779 - - - - - 778  
 |||||  
 DB 1081 KEGPIPPDQNDMSFFKMLFLPESARWIOPTHGKNSLNSGQSPKQVLISGPEKSVEG 1140

QY 779 - - - - - 778  
 |||||  
 DB 1141 QNPLSEKKNVVVGKGEFTKDVGLKEMPPSSKNLFLTLNDLHENTHNOEKKIOEIEK 1200

QY 779 - - - - - 778  
 |||||  
 DB 1201 KETLIOENVLPIQHTVGTKNFMKNLFLSTRQWVEGSGAVAPVQDFRSLNDSTNR 1260

QY 779 - - - - - 778  
 |||||  
 DB 1261 TKKHTAFSKGEBENLBGLAQNOTQOIVEKYACTTRISPNTSOONFYQRSKRALQFRL 1320

QY 779 - - - - - 778  
 |||||  
 DB 1321 PLEETLEKRIIVDTSTQMSKNMKHLPLSTLQIDVNEKKGALIOSPLSDCLTRSHSI 1380

QY 779 - - - - - 778  
 |||||  
 DB 1381 PQANSPLPIAKVSSFPSTIRPIYLRLVLFQDNSSHLPAASYRKKDSQVQESSHFLQAKK 1440

QY 779 - - - - - 778  
 |||||  
 DB 1441 NNLSTALITLEMTGDQREVSLGTSATNSVYKKVENTVLPKPDLPKTSKGVELLPKVI 1500

QY 779 - - - - - 778  
 |||||  
 DB 1501 YQKDLPTIETNSGSPGHLDLVGSLLOGTGKAIKNNENRQKVPFLRVATESSAKTPSK 1560

QY 779 - - - - - 778  
 |||||  
 DB 1561 LLDPLAMDNHGTQIPKEEMKQSEKSPKTAFFKKKDTLLSLNACESNHAIAINBQONK 1620

QY 779 - - - - - 778  
 |||||  
 DB 1621 ELEVWAKOGRTERLCSONPVLKRRHOREITRTTLOSQOEBDDDTISVEMKKEEDPIY 1680

QY 820 DEDENOSPRSFQKTRHYFIAVERLMDYGMSSPHVLRNRAQSGSVQFQKVVFOEFTD 879  
 |||||  
 DB 1681 DEDENOSPRSFQKTRHYFIAVERLMDYGMSSPHVLRNRAQSGSVQFQKVVFOEFTD 1740

QY 880 GSFTQPLXRGELNEHGLGIPYIARAEVDNIMVTRNOASRPSEFSSLSIYEEDORGA 939  
 |||||  
 DB 1741 GSFTQPLXRGELNEHGLGIPYIARAEVDNIMVTRNOASRPSEFSSLSIYEEDORGA 1800

QY 940 EPRKNFVKPNETKTYFMKVYOHMAPTKDEPDKKANAYTSDVDLEKDVHSGILGLYCHT 999  
 |||||  
 DB 1801 EPRKNFVKPNETKTYFMKVYOHMAPTKDEPDKKANAYTSDVDLEKDVHSGILGLYCHT 1860

QY 1000 NTLNPAHGOVTVQEFALFETIETDKSMYFTEENNERCRAPCNTOMDPTEFKENYRHA 1059  
 |||||  
 DB 1861 NTLNPAHGOVTVQEFALFETIETDKSMYFTEENNERCRAPCNTOMDPTEFKENYRHA 1920

QY 1060 INGYIMDTLPGLVMAQORIRIWLTLMSGSENNIHSIHFSGHVFYRKKEEKMALYNLYP 1119  
 |||||  
 DB 1921 INGYIMDTLPGLVMAQORIRIWLTLMSGSENNIHSIHFSGHVFYRKKEEKMALYNLYP 1980

QY 1120 GVETVEMLPSKAGIWRVBCILGELHAGMSTLFLVYSNKCOTPLGMAIGHIRPQITAS 1179  
 |||||  
 DB 1981 GVETVEMLPSKAGIWRVBCILGELHAGMSTLFLVYSNKCOTPLGMAIGHIRPQITAS 2040

QY 1180 GOYGOMAPKLARLHYSGSINAMSTKEPFSWTKYDILAMIIHGIKTQGAROKFSSLYISQ 1239  
 |||||  
 DB 2041 GOYGOMAPKLARLHYSGSINAMSTKEPFSWTKYDILAMIIHGIKTQGAROKFSSLYISQ 2100

|    |      |  |      |
|----|------|--|------|
| Oy | 1240 | FTYMSLDGKKMOYJRNSNGTTLMEFGANDSSGICAHNIENPILARYRLHPHTYSTRS    | 1299 |
| Db | 2101 | FTIMSLDGKKMOYJRNSNGTTLMEFGANDSSGICAHNIENPILARYRLHPHTYSTRS    | 2160 |
| Oy | 1300 | TLRMLKMGCDLNSCSMDLGNESKAISDAQITASSYTTNNFATWSPESKARLHLÖGRSNMR | 1359 |
| Db | 2161 | TLRMLKMGCDLNSCSMDLGNESKAISDAQITASSYTTNNFATWSPESKARLHLÖGRSNMR | 2220 |
| Oy | 1360 | PÖVNNPKEMLOVDÖKTRMKKTYGTTÖGVKSSLTSMYKREPLISSODGHÖMTLFPÖNGKV  | 1419 |
| Db | 2221 | PÖVNNPKEMLOVDÖKTRMKKTYGTTÖGVKSSLTSMYKREPLISSODGHÖMTLFPÖNGKV  | 2280 |
| Oy | 1420 | KVFÖGNDSPFPVYNSLDPPLRLRYLRIRHÖSWVHÖIALRMEVLCEADÖDYL          | 1471 |
| Db | 2281 | KVFÖGNDSPFPVYNSLDPPLRLRYLRIRHÖSWVHÖIALRMEVLCEADÖDYL          | 2332 |

```

RESULT 13
US-08-670-707A-2
Sequence 2, Application US/08670707A
Patent No. 5859204
GENERAL INFORMATION:
APPLICANT: Lollar, John S.
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle Suite 201
City: Boulder
STATE: Colorado
COUNTRY: USA
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/670,707A
FILING DATE: 26-JUN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US94/13200
FILING DATE: 15-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,133
FILING DATE: 11-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Greenlee, Lorraine L.
REGISTRATION NUMBER: 27,894
REFERENCE/DOCKET NUMBER: 75-95F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/499-8080
TELEFAX: 303/499-8089
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2332 amino acids
type: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Liver
US-08-670-707A-2

```

|                       |              |  |            |             |
|-----------------------|--------------|--|------------|-------------|
| Query Match           | 93.0%        | Score 7315   | DB 2       | Length 2332 |
| Best Local Similarity | 62.3%        | Pred. No. 0  |            |             |
| Matches 1452          | Conservative | 0  | Mismatches | 0           |
|                       |              |  | Indels     | 880         |
|                       |              |  | Gaps       |             |
| QY                    | 20           | ATRRRYLGAVELSMWDMOSDGLPVDAPFPPRPVPSFPFNFSVYKRLPVEFVHLFN        | 79         |             |
| Db                    | 1            | ATRRRYLGAVELSMWDMOSDGLPVDAPFPPRPVPSFPFNFSVYKRLPVEFVHLFN        | 60         |             |
| QY                    | 80           | IAKPRPMWGLLGPITQAEVDTVVYITLKNMASHPVSLHVAVSYWKASGEAYDDQTSQ      | 139        |             |
| Db                    | 61           | IAKPRPMWGLLGPITQAEVDTVVYITLKNMASHPVSLHVAVSYWKASGEAYDDQTSQ      | 120        |             |
| QY                    | 140          | REKEDKVPFGSGSHYVWQVLKENGPMASDPLCLTYSTSHVDYKDLNSGLIGALLYCR      | 199        |             |
| Db                    | 121          | REKEDKVPFGSGSHYVWQVLKENGPMASDPLCLTYSTSHVDYKDLNSGLIGALLYCR      | 180        |             |
| QY                    | 200          | EGSLAKKEKQTLHKFTLLFAVFEDEGSMHSFPLKNSLMODRDAASRAMPKMTVGVYAR     | 259        |             |
| Db                    | 181          | EGSLAKKEKQTLHKFTLLFAVFEDEGSMHSFPLKNSLMODRDAASRAMPKMTVGVYAR     | 240        |             |
| QY                    | 260          | SLPGLIGCHARKSVYVHVITGMGTTPEVHSTLEFGHTFVLVNHROASLEISPIITFLAQTL  | 319        |             |
| Db                    | 241          | SLPGLIGCHARKSVYVHVITGMGTTPEVHSTLEFGHTFVLVNHROASLEISPIITFLAQTL  | 300        |             |
| QY                    | 320          | MDLGFPLFCHSHSHOHGMEAYVKVYDSCPPEERPOLRMKNEAEYD DDDLTJDEMDYRF    | 379        |             |
| Db                    | 301          | MDLGFPLFCHSHSHOHGMEAYVKVYDSCPPEERPOLRMKNEAEYD DDDLTJDEMDYRF    | 360        |             |
| QY                    | 380          | DDNSPSFIOIRSVAKKHKHTWVHTYAAEEEMDVAPLVLAADDSSYKQYLNNGPQIRG      | 439        |             |
| Db                    | 361          | DDNSPSFIOIRSVAKKHKHTWVHTYAAEEEMDVAPLVLAADDSSYKQYLNNGPQIRG      | 420        |             |
| QY                    | 440          | RRYKRVKRYMAVYDETFKTEATQIHESGLIGPLLYGEGDTLLIFRKNQASRPYNTYPHGI   | 499        |             |
| Db                    | 421          | RRYKRVKRYMAVYDETFKTEATQIHESGLIGPLLYGEGDTLLIFRKNQASRPYNTYPHGI   | 480        |             |
| QY                    | 500          | TVVRPLYSRRLPKGYAKHLKDPILPGEJLPKYTWIYVEDGPPKSDPRLTRYSSFVAME     | 559        |             |
| Db                    | 481          | TVVRPLYSRRLPKGYAKHLKDPILPGEJLPKYTWIYVEDGPPKSDPRLTRYSSFVAME     | 540        |             |
| QY                    | 560          | RLASGLIGPLLYCKESVDQNGOIMSKRWVILFSVDEKRSWYLTENIORFLPNPAG        | 619        |             |
| Db                    | 541          | RLASGLIGPLLYCKESVDQNGOIMSKRWVILFSVDEKRSWYLTENIORFLPNPAG        | 600        |             |
| QY                    | 620          | VOLLEDEEFGASIMHSINGVDFDLSQVLCHEVAVWYLLISGAOTDFLSVFFSGYTEKH     | 679        |             |
| Db                    | 601          | VOLLEDEEFGASIMHSINGVDFDLSQVLCHEVAVWYLLISGAOTDFLSVFFSGYTEKH     | 660        |             |
| QY                    | 680          | KWVYEDTTLTPPFSGETVFMSENPGLWITLGCNSDFNRGMTALLKYSCDKNTGDYEE      | 739        |             |
| Db                    | 661          | KWVYEDTTLTPPFSGETVFMSENPGLWITLGCNSDFNRGMTALLKYSCDKNTGDYEE      | 720        |             |
| QY                    | 740          | DSYEDISAVLLSKNNAIEPRFSQNSRHDPSTOKOFNAT                         | 778        |             |
| Db                    | 721          | DSYEDISAVLLSKNNAIEPRFSQNSRHDPSTOKOFNAT                         | 780        |             |
| QY                    | 779          |  | 778        |             |
| Db                    | 781          | KIQNVSSDLMMLLNQSPTRPHGLSLSDIQEAKYTFESFDDPSGALDSDNNSLSSEMTNHRPQ | 840        |             |
| QY                    | 779          |  | 778        |             |
| Db                    | 841          | LHNSGDWVFTPESGDLRLNEKLGTTAATTELKLDKFKVSTSNMLISTIPSDNLAAGTDN    | 900        |             |
| QY                    | 779          |  | 778        |             |
| Db                    | 901          | TSSLGPRPMKVHYDSQDLDTTLFGKKSSPLTFESGGLSLSEENNDSKLLSEGLMNSOESSW  | 960        |             |
| QY                    | 779          |  | 778        |             |
| Db                    | 961          | GKNVSTESGRFLFKKRAHGRALLLKQNALFVYSLSLKTNTKSNSNATNKRKTHIDGSL     | 1020       |             |
| QY                    | 779          |  | 778        |             |



Db 1021 LIENSPVWONILESDETEKKVPLIHDRLMLDKNATALRLNHSNKTTSKMMEMVQOK 1080  
QY 779 ----- 778  
Db 1081 KEGPLPOAQNBDMSFFKMLPLPESARMIQTHGKNSLNSGOGPSPKQLVSLGPEKSVEG 1140  
QY 779 ----- 778  
Db 1141 QNFLSEKKNVVVGKEFTKDVGLKEMVFPSSRNFLTLNLDNLHENNTNHQEKIOEIEK 1200  
QY 779 ----- 778  
Db 1201 KETLIQENVVLPQIHVTYGTKNFMKNLFLSTRQVSGSYEGAYAPVLQDFRSLNDSTNR 1260  
QY 779 ----- 778  
Db 1261 TKKHTAHFSKGEENLEGLGNOTQOIVEKACTTRISPNTSQQNVYQSRKALKQFRL 1320  
QY 779 ----- 778  
Db 1321 PLEETELEKRIIVDDTSTOWSKMKHLTPSLTLQIDYNEKEKGAITQSPUSDCLFRSHSI 1380  
QY 779 ----- 778  
Db 1381 POANRSPPLIAKVSSFPSIRPIYLRVLFDQNSSHLPAASVYRKKDQVQESSHFLQGAKK 1440  
QY 779 ----- 778  
Db 1441 NNLSLAILTEMTGDQREVSGISGTSATNSVYTKKVENTVLPKPDLPKTSKVELLPKVIH 1500  
QY 779 ----- 778  
Db 1501 YQKDLFPTETSGSPGHLDLVEGSLLOGTGEGAIKWNNEANRPGKVPFLRVATESAKTPSK 1560  
QY 779 ----- 778  
Db 1561 LLDPLAMDNHYGTQIPKEEMKSQEKSPKTAKKKDTILSLNACSNHAAIAINEGQNKP 1620  
QY 779 ----- 819  
Db 1621 EIEYTWAKOGTERLCSQNPFLKHOIREITRTTLOSQDEEIDYDDTISVEKKEDFDIY 1680  
QY 820 DEDENQSPRSFOKTRHXYFLAAVERLMDYGMSSSPHVLNRRAQSSGVPOFKKVVFOEFTD 879  
Db 1681 DEDENQSPRSFOKTRHXYFLAAVERLMDYGMSSSPHVLNRRAQSSGVPOFKKVVFOEFTD 1740  
QY 880 GSFTOPLRYGELNHLGLGPTIRAEVEDNTIMVTFRNQASRPYSYSSLSIEEDQROGA 939  
Db 1741 GSFTOPLRYGELNHLGLGPTIRAEVEDNTIMVTFRNQASRPYSYSSLSIEEDQROGA 1800  
QY 940 EPRKNFYKPNTEKTYFMKVQHMAAPTDEDFCKAWAYESDVLDLEKDVHSGLIGPLLVCHT 999  
Db 1801 EPRKNFYKPNTEKTYFMKVQHMAAPTDEDFCKAWAYESDVLDLEKDVHSGLIGPLLVCHT 1860  
QY 1000 NTLNPAHGRQVTVQDFALFETTFDETKSMYFETENNERNCRAPCNTOIMEDPTFKENYRFA 1059  
Db 1861 NTLNPAHGRQVTVQDFALFETTFDETKSMYFETENNERNCRAPCNTOIMEDPTFKENYRFA 1920  
QY 1060 INGYIMDTLPGLVMAODORIRMYLLSMGNSNENHSHFSGHVFYRKKEEYKMALYNLYP 1119  
Db 1921 INGYIMDTLPGLVMAODORIRMYLLSMGNSNENHSHFSGHVFYRKKEEYKMALYNLYP 1980  
QY 1120 GVEFTVEMLPSPKAGIWRRECLIGEHLHAGMSTFLVYSNKCOTPLGMAAGHTRDFQITAS 1179  
Db 1981 GVEFTVEMLPSPKAGIWRRECLIGEHLHAGMSTFLVYSNKCOTPLGMAAGHTRDFQITAS 2040  
QY 1180 GQYQOMAPKLARLHYSGSINAMSTKEPFSWIKVDLLAPMIHIGIKTOGAROKFSSLYTSQ 1239  
Db 2041 GQYQOMAPKLARLHYSGSINAMSTKEPFSWIKVDLLAPMIHIGIKTOGAROKFSSLYTSQ 2100  
QY 1240 FIIMYSLDGKKWQTYRGNSGTGLMVFEGNVDSGIGKHINFPPIIARYIRLHPHTYSIRS 1299

Db 2101 FIIMYSLDGKKWQTYRGNSGTGLMVFEGNVDSGIGKHINFPPIIARYIRLHPHTYSIRS 2160  
QY 1300 TLRMELMGCDLNSCAMPJGMEKSAISDAQITPSSYFTNMFPATWSPSKARLHLOGRSNMR 1359  
Db 2161 TLRMELMGCDLNSCAMPJGMEKSAISDAQITPSSYFTNMFPATWSPSKARLHLOGRSNMR 2220  
QY 1360 POVNNPKEMLOVDFQKTKMYGVTTQGVKSLTSMYKREFLISSQDGHQWTLFFQNGKV 1419  
Db 2221 POVNNPKEMLOVDFQKTKMYGVTTQGVKSLTSMYKREFLISSQDGHQWTLFFQNGKV 2280  
QY 1420 KVFQGNQDSFTPVNSLDPPLLTRYLRIHQPSWYHQIALRMEVLGCEKADLY 1471  
Db 2281 KVFQGNQDSFTPVNSLDPPLLTRYLRIHQPSWYHQIALRMEVLGCEKADLY 2332

RESULT 14  
US-09-037-601-2  
; Sequence 2, Application US/09037601  
; Patent No. 6180371  
; GENERAL INFORMATION:  
; APPLICANT: Lollar, John S.  
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Circle Suite 201  
; CITY: Boulder  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/037,601  
; FILING DATE: 26-JUN-1996  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/US94/13200  
; FILING DATE: 15-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/212,133  
; FILING DATE: 11-MAR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/864,004  
; FILING DATE: 07-APR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ferber, Donna M.  
; REGISTRATION NUMBER: 33,878  
; REFERENCE/DOCKET NUMBER: 75-95F  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 303/499-8080  
; TELEFAX: 303/499-8089  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2332 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: protein  
; HYDROTHERICAL: YES  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; TISSUE TYPE: Liver  
US-09-037-601-2

Query Match 93.0%; Score 7315; DB 4; Length 2332;  
Best Local Similarity 62.3%; Pred. No. 0;





QY 1300 TLRMELMCCDLNCSGMPGLGEMSKAISDAQITASSYFTNMFATWSPSKARLHLOGRSNMR 1359  
| | | | |  
Db 2161 TLRMELMCCDLNCSGMPGLGEMSKAISDAQITASSYFTNMFATWSPSKARLHLOGRSNMR 2220  
| | | | |  
QY 1360 POWNPKRMLOVDFOKTKMVTGVTQGVKSLLTSMYVKFELLSSODGOWTLFPONGKV 1419  
| | | | |  
Db 2221 POWNPKRMLOVDFOKTKMVTGVTQGVKSLLTSMYVKFELLSSODGOWTLFPONGKV 2280  
| | | | |  
QY 1420 KVFQGNODSFYVNVNSLDPELLTRYLRHPQSMVHQIALRMEVLGCEADLY 1471  
| | | | |  
Db 2281 KVFQGNODSFYVNVNSLDPELLTRYLRHPQSMVHQIALRMEVLGCEADLY 2332  
| | | | |  
RESULT 15  
PCT-US93-03275-4.  
; Sequence 4, Application PC/TUS9303275  
; GENERAL INFORMATION:  
; APPLICANT: Lollar, John S.  
; APPLICANT: Runge, Marshall S.  
; TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Kilpatrick & Cody  
; STREET: 1100 Peachtree Street, Suite 2800  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: US  
; ZIP: 30309-4530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/03275  
; FILING DATE: 19930407  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/864004  
; FILING DATE: 07-APR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Padst, Patrea L.  
; REGISTRATION NUMBER: 31,284  
; REFERENCE/DOCKET NUMBER: EMO 106PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 404-815-6508  
; TELEFAX: 404-815-6555  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2332 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapien  
; TISSUE TYPE: Liver cDNA sequence  
; PCT-US93-03275-4  
Query Match 93.0%; Score 7315; DB 5; Length 2332;  
Best Local Similarity 62.3%; Pred. No. 0;  
Matches 1452; Conservative 0; Mismatches 0; Indels 880; Gaps 1;  
QY 20 ATRRYLGAVALSDMYMOSDLGELPYDARFPVRPKSPFNTSVVKKTLFEVETVHLFN 79  
| | | | |  
Db 1 ATRRYLGAVALSDMYMOSDLGELPYDARFPVRPKSPFNTSVVKKTLFEVETVHLFN 60  
| | | | |  
QY 80 IAKPPPMGLGPTIOAEVYDTVVITLKNMASHPVSLHAGVSYKASGAEYDDQTSQ 139  
| | | | |

Db 61 IAKPPPMGLGPTIOAEVYDTVVITLKNMASHPVSLHAGVSYKASGAEYDDQTSQ 120  
| | | | |  
QY 140 REKEDDKVFPGSGSHYVWQVLKENGEMASDPLCTYSLSHVDLYKDLNSGLIGALVCR 199  
| | | | |  
Db 121 REKEDDKVFPGSGSHYVWQVLKENGEMASDPLCTYSLSHVDLYKDLNSGLIGALVCR 180  
| | | | |  
QY 200 EGS LAKEKTQTLLKFTLFAVDEGKSHSFTKNSLMODROAASRAAPKMHYNGYNR 259  
| | | | |  
Db 181 EGS LAKEKTQTLLKFTLFAVDEGKSHSFTKNSLMODROAASRAAPKMHYNGYNR 240  
| | | | |  
QY 260 SLPGILGCHRSKYVMHVIGMTTPVHSLFLEGHFTLVNRHQASLESTPTFLTAOTLL 319  
| | | | |  
Db 241 SLPGILGCHRSKYVMHVIGMTTPVHSLFLEGHFTLVNRHQASLESTPTFLTAOTLL 300  
| | | | |  
QY 320 MDLGFLEFCHISHQHDMGEAYVYKVDSCPEEPOLRMKNNEAEEDYDDLTJDEMDVVR 379  
| | | | |  
Db 301 MDLGFLEFCHISHQHDMGEAYVYKVDSCPEEPOLRMKNNEAEEDYDDLTJDEMDVVR 360  
| | | | |  
QY 380 DDNSPSFTQIRSVAKKHPKTVWHYIAEEDMDYAPLVLPADDRSYKSYTLNNGPORIG 439  
| | | | |  
Db 361 DDNSPSFTQIRSVAKKHPKTVWHYIAEEDMDYAPLVLPADDRSYKSYTLNNGPORIG 420  
| | | | |  
QY 440 RYKKYRPMAYTDEFTKTRREALTOHESGILGPLYGEVGDTLILFKNQSAPRYNTPHGI 499  
| | | | |  
Db 421 RYKKYRPMAYTDEFTKTRREALTOHESGILGPLYGEVGDTLILFKNQSAPRYNTPHGI 480  
| | | | |  
QY 500 TDVRLPYSRRLKPGVYHLKDFILPGELFKYKWTYVEDGPTKSDPRLTRYSSFYVME 559  
| | | | |  
Db 481 TDVRLPYSRRLKPGVYHLKDFILPGELFKYKWTYVEDGPTKSDPRLTRYSSFYVME 540  
| | | | |  
QY 560 RDLASGLIPLLICYKESVDQGNQIMSDKNVILFSVDEKRSWLTENTQRLPNPAG 619  
| | | | |  
Db 541 RDLASGLIPLLICYKESVDQGNQIMSDKNVILFSVDEKRSWLTENTQRLPNPAG 600  
| | | | |  
QY 620 VOLEDEPQASNMHSINGVDFDSLOLSVCLHEVYWTLSGAOTDPLSYFFSGYTFKH 679  
| | | | |  
Db 601 VOLEDEPQASNMHSINGVDFDSLOLSVCLHEVYWTLSGAOTDPLSYFFSGYTFKH 660  
| | | | |  
QY 680 KMYEDTLTLFPFSGETVMSMENPGLWILGCHNSDFRNKGTALLKYSSCDKNTGDYE 739  
| | | | |  
Db 661 KMYEDTLTLFPFSGETVMSMENPGLWILGCHNSDFRNKGTALLKYSSCDKNTGDYE 720  
| | | | |  
QY 740 DSYEDISATLKSNNAIERFSQNSRHPSTROKOPNAI 778  
| | | | |  
Db 721 DSYEDISATLKSNNAIERFSQNSRHPSTROKOPNAI 780  
| | | | |  
QY 779 778  
| | | | |  
Db 781 KIONVSSDPLMLLROSPTPHGLSLSDLOEAKYETFPDDPSGALDSNNSISEMTHFRPQ 840  
| | | | |  
QY 779 778  
| | | | |  
Db 841 LHHSGDMVFTPESGLQLRLNEKLGTTAATLKKLDFKVSSTSNLSTIPSDNLAAGTDN 900  
| | | | |  
QY 779 778  
| | | | |  
Db 901 TSLGPPSMVHYVDSQDTTLFGKKSSPLTESGCLSLSEENNDKLLSEGLMNSOESSW 960  
| | | | |  
QY 779 778  
| | | | |  
Db 961 GKNVSTESGRLEFKGRAHPALLTKDNALFKVVISLTKTKTSNNSATNRKTHIDGPSL 1020  
| | | | |  
QY 779 778  
| | | | |  
Db 1021 LIENSPVQONLLESDTEKKVTPLIHDMLMDKNATLRLHMSNKTTSKKNMEMYQOK 1080  
| | | | |  
QY 779 778  
| | | | |  
Db 1081 KEGPIPPAQNDPMSEFFKMLFLPESARWIOPTHGRKNSLNSGQSPKOLVSLGPEKSYEG 1140  
| | | | |  
QY 779 778  
| | | | |  
Db 1141 QNFLESKRNVVVGKGEFTKDVGLKEVFPSSRNLFVTLNLDLHENTHNOEKKIOEETEK 1200  
| | | | |

QY 779 ----- 778  
Db 1201 KETLIGENVLPQIHFTVGTKNEMKNLFLSTRQNVESYEGAYAPVLODERSLNDSTNR 1260  
QY 779 ----- 778  
Db 1261 TKKHTAHSKKGEEENLEGLOQTQOIVEKXACTTIRISPTNSQONFVYORSKRALKOPRL 1320  
QY 779 ----- 778  
Db 1321 PLEETELEKRIIVDDTSTQWSKNMKHLTPSTLQIDYNEKEGAIQSPLSDCLTRSHSI 1380  
QY 779 ----- 778  
Db 1381 PQANRSPPIAKVSSPSIRPIYLRVLFODNSSHLPAASYRKKDSGVQESSHFLQAKK 1440  
QY 779 ----- 778  
Db 1441 NNLSLAILTEMTGDQREVSLGTSATNSVTYKKVENTVLPKPLPKTSKVELLPKVHI 1500  
QY 779 ----- 778  
Db 1501 YQKDLFPTETNSGSPGHLDLVEGSLQGTGEGAIKWEANRPCKVPELRVATESSAKTPSK 1560  
QY 779 ----- 778  
Db 1561 LLDPLAMDNHYGTQIPKREMKSQEKSPKTAFAKKDTILSLNACESNHAIAINEGONKP 1620  
QY 779 -----PVLRKHOREIRRTTLOSDQDEIDYDDTISVEMKKEDFDIY 819  
Db 1621 ELEVWAKQGRTERLCSQNPVLRKHQREIRRTTLOSDQDEIDYDDTISVEMKKEDFDIY 1680  
QY 820 DEDENOSRSPFOKTRRHFTIAVERLMDYGMSSSPHVLNRNAGSGSVQPKKVPQOFTD 879  
Db 1681 DEDENOSRSPFOKTRRHFTIAVERLMDYGMSSSPHVLNRNAGSGSVQPKKVPQOFTD 1740  
QY 880 GSFTQPLRGELNEHILGLPYIRAEVDNIMVTFRNQASRPYSFSSLSIYEEDOROGA 939  
Db 1741 GSFTQPLRGELNEHILGLPYIRAEVDNIMVTFRNQASRPYSFSSLSIYEEDOROGA 1800  
QY 940 EPRKNFVAPNETKTYFWKVQHMAPTKDEFDCKAMAYFSDVLEKDVHSGILGPLVCHT 999  
Db 1801 EPRKNFVAPNETKTYFWKVQHMAPTKDEFDCKAMAYFSDVLEKDVHSGILGPLVCHT 1860  
QY 1000 NTLNPAHGRQVTOEFALEFTIPEDTKSWTYTENMERCRAPCNIOMEDPTEKENYRFHA 1059  
Db 1861 NTLNPAHGRQVTOEFALEFTIPEDTKSWTYTENMERCRAPCNIOMEDPTEKENYRFHA 1920  
QY 1060 INGYIMDTLPGLVAAODQIRIMYLYLSMGSNENIHSIHSGHVFYVRKKEEYKMALYNLYP 1119  
Db 1921 INGYIMDTLPGLVAAODQIRIMYLYLSMGSNENIHSIHSGHVFYVRKKEEYKMALYNLYP 1980  
QY 1120 GVFEVEMLPKAGIWRVECLIGELHLAGMSTLFLVYSNKCOTPLGMAAGHIRDOFITAS 1179  
Db 1981 GVFEVEMLPKAGIWRVECLIGELHLAGMSTLFLVYSNKCOTPLGMAAGHIRDOFITAS 2040  
QY 1180 GQYOGMAPKLARLHYSGSINAMSTKEPESWIKVDLLAPMIHGIKTQAGAROKFSSLYISQ 1239  
Db 2041 GQYOGMAPKLARLHYSGSINAMSTKEPESWIKVDLLAPMIHGIKTQAGAROKFSSLYISQ 2100  
QY 1240 FIIMYSLDGKKMOTYRGNSTGTLMVFFGNVDSGIIKHNIENPILARIYIRLPHYHSIRS 1299  
Db 2101 FIIMYSLDGKKMOTYRGNSTGTLMVFFGNVDSGIIKHNIENPILARIYIRLPHYHSIRS 2160  
QY 1300 TLRRELKMGCDINSCSMPLGMSKASIDAQITASSYFTNMFATWSPSKARLHLQGRSNAMR 1359  
Db 2161 TLRRELKMGCDINSCSMPLGMSKASIDAQITASSYFTNMFATWSPSKARLHLQGRSNAMR 2220  
QY 1360 POVNNPKEMLOVDFOKTMKVGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFONGKV 1419  
Db 2221 POVNNPKEMLOVDFOKTMKVGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFONGKV 2280

QY 1420 KVFOGNQDSFTFPVYVNSLDPPLTRYLRIHPQSWVHQIALRMEVLCGEAODLY 1471  
Db 2281 KVFOGNQDSFTFPVYVNSLDPPLTRYLRIHPQSWVHQIALRMEVLCGEAODLY 2332

Search completed: November 17, 2001, 13:24:06  
Job time: 115 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2001, 04:58:30 : Search time 389.82 Seconds  
(without alignments)  
2768.374 Million cell updates/sec

Title: US-09-689-430-1\_COPY\_150\_4914

Perfect score: 4765

Sequence: 1 cttcttcaagtaacagta.....ttgggtcgttcgttcgagtc 4765

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapept 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA:\*  
1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/PCRDUS\_COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID               | Description        |
|------------|--------|-------------|--------|---------------------|--------------------|
| 1          | 4413   | 92.6        | 4629   | 2 US-08-484-891-7   | Sequence 7, Appl1  |
| 2          | 4411.6 | 92.6        | 9354   | 1 US-08-683-8398-2  | Sequence 2, Appl1  |
| 3          | 4357.6 | 91.5        | 4670   | 3 US-08-717-294-41  | Sequence 41, Appl1 |
| 4          | 4330.8 | 90.9        | 4999   | 4 US-09-470-618-14  | Sequence 14, Appl1 |
| 5          | 4330.8 | 90.9        | 4999   | 4 US-09-364-862-14  | Sequence 14, Appl1 |
| 6          | 4259.4 | 89.4        | 11933  | 4 US-09-470-618-13  | Sequence 13, Appl1 |
| 7          | 4259.4 | 89.4        | 11933  | 4 US-09-364-862-13  | Sequence 13, Appl1 |
| 8          | 3748   | 78.7        | 5035   | 2 US-08-882-083-1   | Sequence 1, Appl1  |
| 9          | 3748   | 78.7        | 5035   | 2 US-08-558-107-1   | Sequence 1, Appl1  |
| 10         | 3748   | 78.7        | 5035   | 3 US-09-243-539-1   | Sequence 1, Appl1  |
| 11         | 3189.8 | 66.9        | 4334   | 2 US-08-670-707A-38 | Sequence 38, Appl1 |
| 12         | 3189.8 | 66.9        | 4334   | 4 US-09-037-601-38  | Sequence 38, Appl1 |
| 13         | 2605.4 | 54.7        | 4451   | 4 US-08-717-294-42  | Sequence 42, Appl1 |
| 14         | 2336.4 | 49.0        | 9009   | 1 US-07-864-0048-3  | Sequence 3, Appl1  |
| 15         | 2336.4 | 49.0        | 9009   | 1 US-08-251-937A-3  | Sequence 3, Appl1  |
| 16         | 2336.4 | 49.0        | 9009   | 1 US-08-212-133A-1  | Sequence 1, Appl1  |
| 17         | 2336.4 | 49.0        | 9009   | 1 US-08-474-503-1   | Sequence 1, Appl1  |
| 18         | 2336.4 | 49.0        | 9009   | 2 US-08-670-707A-1  | Sequence 1, Appl1  |
| 19         | 2336.4 | 49.0        | 9009   | 4 US-09-037-601-1   | Sequence 1, Appl1  |
| 20         | 2336.4 | 49.0        | 9009   | 5 PCT-US93-03275-3  | Sequence 3, Appl1  |
| 21         | 2336.4 | 49.0        | 9009   | 5 PCT-US94-13200-1  | Sequence 1, Appl1  |
| 22         | 2335.4 | 49.0        | 7056   | 1 US-08-121-202-1   | Sequence 1, Appl1  |
| 23         | 2334.8 | 49.0        | 8241   | 6 5171844-1         | Patent No. 5171844 |
| 24         | 2233.2 | 49.0        | 8967   | 1 US-08-366-851A-1  | Sequence 1, Appl1  |
| 25         | 2275.2 | 47.7        | 6999   | 1 US-08-276-594A-1  | Sequence 1, Appl1  |
| 26         | 1848   | 38.8        | 7032   | 4 US-09-324-867-1   | Sequence 1, Appl1  |
| 27         | 1802.8 | 37.8        | 7493   | 1 US-08-212-133A-7  | Sequence 7, Appl1  |

|    |        |      |      |                     |                    |
|----|--------|------|------|---------------------|--------------------|
| 28 | 1802.8 | 37.8 | 7493 | 1 US-08-474-503-5   | Sequence 5, Appl1  |
| 29 | 1802.8 | 37.8 | 7493 | 2 US-08-670-707A-5  | Sequence 5, Appl1  |
| 30 | 1802.8 | 37.8 | 7493 | 4 US-09-037-601-5   | Sequence 5, Appl1  |
| 31 | 1802.8 | 37.8 | 7493 | 5 PCT-US94-13200-5  | Sequence 5, Appl1  |
| 32 | 1751.2 | 36.8 | 6402 | 2 US-08-670-707A-36 | Sequence 36, Appl1 |
| 33 | 1751.2 | 36.8 | 6402 | 4 US-09-037-601-36  | Sequence 36, Appl1 |
| 34 | 1240.4 | 26.0 | 1623 | 1 US-08-121-202-3   | Sequence 3, Appl1  |
| 35 | 893.4  | 18.7 | 1130 | 1 US-07-864-0048-1  | Sequence 1, Appl1  |
| 36 | 893.4  | 18.7 | 1130 | 1 US-08-251-937A-1  | Sequence 1, Appl1  |
| 37 | 893.4  | 18.7 | 1130 | 1 US-08-212-133A-5  | Sequence 1, Appl1  |
| 38 | 893.4  | 18.7 | 1130 | 1 US-08-474-503-3   | Sequence 3, Appl1  |
| 39 | 893.4  | 18.7 | 1130 | 2 US-08-670-707A-3  | Sequence 3, Appl1  |
| 40 | 893.4  | 18.7 | 1130 | 4 US-09-037-601-3   | Sequence 3, Appl1  |
| 41 | 893.4  | 18.7 | 1130 | 5 PCT-US93-03275-1  | Sequence 1, Appl1  |
| 42 | 893.4  | 18.7 | 1130 | 5 PCT-US94-13200-3  | Sequence 3, Appl1  |
| 43 | 301.8  | 6.3  | 6909 | 2 US-08-804-196-1   | Sequence 1, Appl1  |
| 44 | 301.8  | 6.3  | 6909 | 2 US-08-658-340-1   | Sequence 1, Appl1  |
| 45 | 301.8  | 6.3  | 6909 | 3 US-08-746-111-26  | Sequence 26, Appl1 |

## ALIGNMENTS

RESULT 1  
US-08-484-891-7  
Sequence 7, Application US/08484891  
Patent No. 5935935  
GENERAL INFORMATION:  
APPLICANT: Connolly, Sheila  
APPLICANT: Kaleko, Michael  
APPLICANT: Smith, Theodore  
TITLE OF INVENTION: Adenoviral Vectors for  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cecchi, Stewart & Olstein  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,891  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/218,335  
FILING DATE: 25-MAR-1994  
APPLICATION NUMBER: 08/074,920  
FILING DATE: 10-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Olstein, Elliot M.  
REGISTRATION NUMBER: 24,025  
REFERENCE/DOCKET NUMBER: 271010-273  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4629 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA primer  
FEATURE:  
NAME/KEY: Factor VIII cDNA with  
NAME/KEY: B domain deleted



|    |      |  |      |
|----|------|--|------|
| OY | 2371 | atggaaaaccaggctatgatactctcggggtgccacaactcagactcttcggaacaagagc    | 2430 |
| Db | 2101 | ATGGAAAACCCAGGTCTATGTGATTCGGGGGTCCCAACCTCAGACTTTCGGAAACGAGGC     | 2160 |
| OY | 2431 | atgacgccttactgaaggtcttcctagcttgtagcaagaacacgtggtatattacgaagac    | 2490 |
| Db | 2161 | ATGACCCCTTACTGANGGTTTCTACTTGTGCAAGAACCTGGTGATTTATTCGAGGAC        | 2220 |
| OY | 2491 | agttatgaagatacttcagatatactctgcgagtaaaacaatctccatctgaaaccaagaagc  | 2550 |
| Db | 2221 | AGTTATGAACATATTTCAGATACTCTGCGTAGTAAAAACAATGGCATTTGAACCAAGAAGC    | 2280 |
| OY | 2551 | ttctccagaacttcaagaacacccctagcacttaggcgaagaacattaaatgcaccccaca    | 2610 |
| Db | 2281 | TTCTCCAGATTTCAAGACACCCCTACCTAGGCAAAAAGCAATTTAAATGCCACCCACACA     | 2340 |
| OY | 2611 | gtcttgaagaagccataaagggaataactctgtactactcttcgtacgtatcaagaagaa     | 2670 |
| Db | 2341 | GTCTTGAAGCCATCAACGGGAATACTCGTACTCTTCACTCACTGATCAAGAGGAA          | 2400 |
| OY | 2671 | atctacatgatataccatatcagttctgaacttgaagaagaagatcttgcataattcatgat   | 2730 |
| Db | 2401 | ATTGACATATGATATACCATATACGTATGTAATGAATGAAGAAGATTTGCACATTATGAT     | 2460 |
| OY | 2731 | gaagatgaaaatcaagagcccccgagctcttcaaaagaanaacgacactattatgtct       | 2790 |
| Db | 2461 | GAGATGAAAATCAAGACCCCCGACCTTTCAAAAGAAACACGACACTATTTATTGCT         | 2520 |
| OY | 2791 | gcagtggaagagctctcgggattatgtagatgaatgaactcccaacatgtctctaagaacaag  | 2850 |
| Db | 2521 | GCACTGGAGAGGCTCTCGGGATTATGGGATGAGTACTCCCAACATGTTCTTAAGAACAGG     | 2580 |
| OY | 2851 | gtcaagatgagagtgctccctcagattccaagaagcttgctttccagaagaaattacgtatgac | 2910 |
| Db | 2581 | GCTAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTTTCAGGAATTTACTGATGGC       | 2640 |
| OY | 2911 | tccttctcagccctataacgltggaagaactaaatgaacatlttgagactcctcgtggcca    | 2970 |
| Db | 2641 | TCCTTTACTCAGCCCTTATACGCTGGAAGAACTTAATGAACATTTGGGACCTCTGGGGCCA    | 2700 |
| OY | 2971 | tataaagacgaagatgaaagataatacatcgttaacttccagaaatcaagcctctcgt       | 3030 |
| Db | 2701 | TATATAAGACCAACAAAGTTGAAGATATATCAAGTACCTTCAGAAATCAAGCCCTCCGT      | 2760 |
| OY | 3031 | cccatctcccttcatctcagccttattcttctatgtaggaagtcgaaggcaagagcgaa      | 3090 |
| Db | 2761 | CCCATTCCTTCTATTCCTAGCCTTATTTCTATAGGAAGATCAAGGCAAGSGACGAA         | 2820 |
| OY | 3091 | cctagaanaaactctgcaagccctaaatgaacaaaacttactcttggaaagtgcacat       | 3150 |
| Db | 2821 | CCTAGAAAAAAGCTTTGCAAGCCTATATGAACCAAAAGCTACTTTTGGAAAGTGCACAT      | 2880 |
| OY | 3151 | catatgaccccaataaagatgaatgttgactctgcgcaaaagccgtggctattctctgatgtc  | 3210 |
| Db | 2881 | CATATGGCACCCACTAAAGATGAGTTGACTCTCAAGAGCCTGGGCTTATATTCCTCGATGTT   | 2940 |
| OY | 3211 | gaacttgaaaaaagtgtgactcgaagccgatctggaacccctctcgtctgtcgcaactaac    | 3270 |
| Db | 2941 | GACCTGAAAAAATGTGCACTCAGGCTGATTTGAACCCCTTCTGTGGTGTGCACACTAAC      | 3000 |
| OY | 3271 | acactgaaccctctcatctgggagacaagtgcgaagtagcagagatctgcctcttctaac     | 3330 |
| Db | 3001 | ACACTGAACCTCTCATGGGAGACAAAGTACAGTACAGTAATTTGCTTTTCACC            | 3060 |
| OY | 3331 | atctcttgatgaagacaaagctggtactctcacttgaanaataatgaaagaacatcgaggct   | 3390 |
| Db | 3061 | ATCTTTGATAGACCAAAAGCTGTGACTTCACTGGAATAATAGGAAGCAAACTGCAGGGCT     | 3120 |
| OY | 3391 | ccctgcaatctcaagttgaaagatcccaatttaagagaattatcgttcctatgcacatc      | 3450 |
| Db | 3121 | CCCTGCATATCCAGATGGAAGATCCCACTTTTAAAGGAATATATCGTTTCCATGCATC       | 3180 |

|    |      |  |      |
|----|------|--|------|
| QY | 3451 | atggtcacataatggataacacaccggcttagtaatggtcccaagatccaagaatttga        | 3510 |
| Dp | 3181 | AATGGCTACATAAATGAAGTATACACTGGCTTAAGTAAATGGCTACAGATCAAAAGATTTGCA    | 3240 |
| QY | 3511 | tgatcatcgtcccaagatgggcaagcaatgaacaatcatcttcatcatcttaagtgagcat      | 3570 |
| Dp | 3241 | TGGTATCTGGCTCAGCATGGGCAGCATGAAGAACATCTCATTTTCATTTCAGTGGACAT        | 3300 |
| QY | 3571 | gtgtcactcgttacgnaaaaaaaggaggtataaatlgygacgttacaatctcatccaggt       | 3630 |
| Dp | 3301 | GTTGTTCACTGTACGAAAAAAGAGGAGTAAATAATGGCACTGTACAAATCTGATCCAGGT       | 3360 |
| QY | 3631 | gtttttgaagacgtlgyaaatcgttaccatccaagaagcttgyaaatttggcgggttgaatgcctt | 3690 |
| Dp | 3361 | GTTTGTGAGACAGTGGAAATGTTAACCATCCAAAGCTGGAATTTGGCGGGTGGAAATGCCCTT    | 3420 |
| QY | 3691 | atggcgaagcatccaatcgtctggatgtagaagaaccttttcgtgttgaagcaataagtt       | 3750 |
| Dp | 3421 | ATTTGGCAGCACTTACATAGCTGGAGATGAGCACCTTTTCTGTGTGATGACGCAATTAAGTT     | 3480 |
| QY | 3751 | cagaatcccccctgggnaatlygctcttcggaacatatagaatcttcaagatlacagcttcaaga  | 3810 |
| Dp | 3481 | CAGACTCCCCGGGAATAGGCTTCTGGACACATTAAGATTTTCAGATTACAGCTTCAGGA        | 3540 |
| QY | 3811 | caataatggaacaglyggcccccaaaagtctggccagaatcattatccggtacaatcaatgc     | 3870 |
| Dp | 3541 | CAATATGTGACAGTGGGCCCCCAAAAGCTGGGCCAGACTTCATTAATTCGGATCAATCAATGCC   | 3600 |
| QY | 3871 | tgagagaccacaaggagacctttcttggatcaagglygattctgttggaccacaatgatttt     | 3930 |
| Dp | 3601 | TGGAGCACCAAGGAGCCCTTTTCTTGGATCAAGTGTGATCTGTTGGCACCAATGATTAATTT     | 3660 |
| QY | 3931 | caagcactcaagaagcccaagggtggccgtccgaagtctccagcctcaaatctcgaatt        | 3990 |
| Dp | 3661 | CAGGCAATCAAGAACCCAGGGTGGCCCGTCAGAAATTTTCCAGGCTTACATCTCTGAGTTT      | 3720 |
| QY | 3991 | atcatcatgatatagctcttgaatggnaaaglyggcagacttalcgaagaattccacgtga      | 4050 |
| Dp | 3721 | ATCATCATGTATATAGTCTTGAATGGGAAGATGGCAGACACTTATCGAGGAATTCCTCGGA      | 3780 |
| QY | 4051 | accttaatggtctctcttggcaaltgfygattatctcgggataataaacacaatatctttac     | 4110 |
| Dp | 3781 | ACCTTAATGGCTCTTCTTTGGCAATGTGGATTCATCTGGGATTAACAAATATTTTATTAAC      | 3840 |
| QY | 4111 | ccccaatatttgcgtgatcatcatccgtttggaccacaacatcatatagaatccagagact      | 4170 |
| Dp | 3841 | CCCTCCATTATTTGCTGCATATCATCCGTTTGGACCCACACATTAATTAAGCATTTGGCAGCACT  | 3900 |
| QY | 4171 | cttcgcactlgyagttgatalggtctgataltaaataglttgcaagatgcatctlgyaaatgag   | 4230 |
| Dp | 3901 | CTTCCGATGGAGTTGATGGGCTGTGATTAATAATAGTTCAGCATGCCATTTGGGAATGGAG      | 3960 |
| QY | 4231 | agtaagaacataatcagatgtcacagatatctgcttcatcctcatcttacaatatglttggc     | 4290 |
| Dp | 3961 | AGTAAGAACATATTCAGATGTGACAGATTACTGCTTCACTTATTCACAAATATGTTTGGC       | 4020 |
| QY | 4291 | acctgggtctcccttcaaaaagctlgaacttcaactccaaggaggaagatgacgtccggagagact | 4350 |
| Dp | 4021 | ACCTGGGTCTCTCTTCAAAAAGCTCCACATTTCAACTTCMAAGGAGAGTAATAGCTCGAGAGACTT | 4080 |
| QY | 4351 | caggltgaataatccaaaaggatlygctlgycaagtlygaatccagaagaacaaatgaagtcca   | 4410 |
| Dp | 4081 | CAGGTGAATTAATCCAAAGAGTGGCTGCAAGTGGAGACTTCCAGAAAGACATGAAGTACAA      | 4140 |
| QY | 4411 | ggagtaactactcaaggaggttaaaatctctgcttaccagcatgtatgtyaaaggagttcttc    | 4470 |
| Dp | 4141 | GGAGTAACACTACTCAGGAGAGTAATAATCTTGCTTACAGCATTAATGATAAGAGATTTCCCTC   | 4200 |
| QY | 4471 | atcccgaagagttcaaggltggccaatgaggacctcttttccaatatgynaagaagtlaag      | 4530 |
| Dp | 4201 | ATTTCCAGCACTCAAGATGGCCATCATAGTGGACTCTTTTTCACAAATATGGCAAAAGTAAG     | 4260 |
| QY | 4531 | gttttccaaggaaatcaagaactccttcacaacgtlgtlgtlaaactctcttgcaccaacgttca  | 4590 |

|||||  
Db 4261 GTTTTTCAGGGAATTCAGACTCTTCACACCTGTGTGAACCTCTAGACCCACGGTTA 4320  
QY 4591 ctgactcgtcactcctcgatccaacccccagagttgtgtgacacagatgaccttgagatg 4650  
|||||  
Db 4321 CTGACTCGCTACTCTTCGATTTCACCCCAAGAGTTGGGTGCACCAATTGCGCTGAGATG 4380  
QY 4651 gaagttctggtgctgagagacacagagacctctactgactcgagcgatgc 4699  
|||||  
Db 4381 GAGTGTCTGGGCTGAGGACACAGAGACCTCTACTGAGGGTGGCCACTGC 4429

RESULT 2  
US-08-683-839B-2  
; Sequence 2, Application us/08683839B  
; Patent No. 5744326  
; GENERAL INFORMATION:  
; APPLICANT: ILL. Charles . R. et al.  
; TITLE OF INVENTION: Use of Viral Cis-Acting Post-Transcriptional  
; TITLE OF INVENTION: Regulatory Sequences To Increase Expression of  
; TITLE OF INVENTION: Intronsless Genes Containing Near-Consensus Splice Sites  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHYE & COCKFIELD  
; STREET: 60 State Street, suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/683,839B  
; FILING DATE: 11-MARCH-1996  
; CLASSIFICATION: 435  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Remillaard, Jane E.  
; REGISTRATION NUMBER: 38,872  
; REFERENCE/DOCKET NUMBER: TTI-138  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7440  
; TELEFAX: (617)227-5941  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9354 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 2965..7378  
; US-08-683-839B-2

Query Match 92.6%; Score 4411.6; DB 1; Length 9354;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 4420; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 266 ccacatgcaaatagagctcgcacctgcgtctcttctgtgaccttgcgattcgcttca 325  
|||||  
Db 2960 CCACCTGGTTTATGACTCTCCACCTGCTCTTCTGTGTGCTTTTGGATTCTGCTTTA 3019  
QY 326 gtgcacacagaagatactactggtgcaatgagacgtgtcagacataatgaagaatg 385  
|||||  
Db 3020 GTGCCACCAAGATCTACTGAGGTGAGTGAAGTGTATGCAAAAGTG 3079

QY 386 atctcgtgagctgctcgttgagcgaagattctcctcctagagtgccaaaaatcttccat 445  
|||||  
Db 3080 ATCTCGTAGAGCTGCCTGTGGAGCGCAAGATTTCTCTAGAGTGCAAAATCTTTTCCAT 3139  
QY 446 tcaaacactcagtcgtgtacaaaaaagactcgttltgtgaattcaacggttacccttca 505  
|||||  
Db 3140 TCAACACCTCACTCTGTATCAAAAAGACTCTGTGTGTGAAATTCACGGTTACCTTTTCA 3199  
QY 506 acatcgtcaagccaagcgccacctgagtcgtgtcgtatagtcctaccatccagcgtag 565  
Db 3200 ACATGCTAAGCCAGGCGCACCGCCACCTGGATGGGTCTGTAGTCTCTACCAATCCAGGCTGAGG 3259  
QY 566 ttatgatacagtggtgcatlaaacacttaagaacaatggtcttccactcgtgactcatg 625  
Db 3260 TTTATGATACAGATGCTGCTATTCACACTTAAGAACATGGCTTCCATCTGTGAGTCTTCATG 3319  
QY 626 ctgttggtgtatccctacacgggaagcttcttgaggagcgtgaatagatgaacccagtc 685  
Db 3320 CTGTTGGTATCCCTACGGAAGAGCTTCTGAGGGAGCTGAATATGATGATCAGACAGTC 3379  
QY 686 aaagagagaagaagatataaagcttcctcctggtgtgaagcatatagtctgagcag 745  
Db 3380 AAAGGAGAAAGACATGATATAAGTCTTCCCTGGTGGAGGCCATACATATGCTGCGCAGG 3439  
QY 746 tcttgaagaagaatggttccaaatggtccttgacccactggtcctactactatatt 805  
Db 3440 TCCTGAAGAGAAATGCTCCATATGGCTGTGACCCACTGTGCTTACCTACATATCTTT 3499  
QY 806 ctcatgtgacctggttaaaagacttgaattcaggtcccatcttgagccctactagtagta 865  
Db 3500 CTCATGTGACCTGCTGTAAGAAACCTTGAATCAGGCTTCATTGGAGCCCTCTACTATATGTA 3559  
QY 866 gagaagggagctcgtgccaaggaagaagacagacctgtgccaattatatactactttg 925  
Db 3560 GAGAAAGGAGTCTGGCCACAGGAAAAAGACACAGACTTGTCACAAATTTATCTACTTTTG 3619  
QY 926 ctgtattgtatgaagggaaaagtctggtcactcagaacaagaactccttgatgagagata 985  
Db 3620 CTGTATTTGATGAAGGAGGAAAGTTGGCACTCAGAAACAAAGAACTCTGTATGAGAGATA 3679  
QY 986 gggatgctgacttgcgtcgtggccctggcctaaatgacacagatgaatggttagtaaca 1045  
Db 3680 GGGATGCTGCATCTGCTGGGCTGGCCCTGCTAAAGACACAGTCAATGATGTATAACA 3739  
QY 1046 ggtctcgtccaggtctgattgtgatacacaagaagaatcaagctatcttgagcagtgtgaa 1105  
Db 3740 GGTCTCTCCAGAGTGTGATTGTGATGCCACAGAAATCATGTTATGGCATGTATTTGAA 3799  
QY 1106 tgggacacacactcctgaagtgcaactcaatatctcctcgaaggtcacacattccttgaaga 1165  
Db 3800 TGGGCACCACTCTGMAAGTGCACATCAATATTCCTGAAAGTCAACATTTCTGTGAGGA 3859  
QY 1166 accatcgcgaagcgcttcccttggaatatctgcgaataacttctcctactcgtccaactctc 1225  
Db 3860 ACCATCGCCAGGCGCTCTGGAAATCTCGCCCAATTAACCTTCTCTTACGTCAACACATCT 3919  
QY 1226 tgaatgaccttggagacgttctactggttgtgtcatatcctctccacaacatgaatgga 1285  
Db 3920 TGATGAGACCTTGAGACAGTTTCTACTGTTTTGTCTATCTCTTCCACCAACATGATGGA 3979  
QY 1286 tgaagactatgtcaaaatagacagctgtccagagaagaccccaactcagaatgaaaaata 1345  
Db 3980 TGAAGGCTTATGTCAAAAGTAGACAGCTGTCCAGAGAAACCCCACTACGATGAAAAATA 4039  
QY 1346 atgaagaagcggaagactatgatagtatcttactgattcgaatgaaatgaaatgtgtcaagt 1405  
Db 4040 ATGAAGAAGCGGAAGACTATGATGATGATCTTACTGATTCGAAATGAGATGTGTGAGGT 4099  
QY 1406 ttgatgtgacaactcctcctcttcaatcgaatctgctcgaatgtgcaaaagaacatccca 1465  
Db 4100 TTGATGATGACAACTCTCTCTCTTTTATCCAAATTCCTCTCAGTTGGCCAAAGACATCTTA 4159  
QY 1466 aaacttgggtacattacatgtcgtcgaagaagagagacttggagatagtcctccttagtcc 1525



Db 4160 AACCTGGGTACATTACTTCTGCTGAGAGAGAGGACTGGAGCTATGCTCCCTTAGTCC 4219  
QY 1526 tcgcccccgatgacagaagtataaagtaataattgacaatrgccctcagcgatgtg 1585  
Db 4220 TCGCCCCGATGACAGAAAGTTATAAAGTCAATATTTGAAACAATGGCCCTCAGGGGATGG 4279  
QY 1586 gtagaagtacaaaaaagctccgattatgatacacagaatgaaccttaagactcgtg 1645  
Db 4280 GTAGGAAGTACAAAAAGTCCGATTTATGGCATACACGATGAAGACCTTTAGACTCGTG 4339  
QY 1646 aagctatcagcagtaacccaagaaactcgtggacccttaacttaacggtggaagtgagaca 1705  
Db 4340 AAGCTATTACACATGAAATCAGGAATCTTGGACCTTTACTTTATGGGGAAGTGGAGACA 4399  
QY 1706 cactgttgatataatttaagaatacaagcagacagacataaacttaacctcagga 1785  
Db 4400 CACTGTTGATTATATTTAAGATCAAGCAACGACCATTAACATCTACCCCTCAGCGAA 4459  
QY 1766 tcactgagatgcgcctccttgatctcaaggagatcaccaaaagtgtaaaacattgaag 1825  
Db 4460 TCACCTGATGTCCTGCTTGTGATTCAAGAGATTAACAAAAGGTGTAAACATTGTGAAG 4519  
QY 1826 atttccaatctgcagagagaaataticaaataaaatggacagttagctagaagatg 1885  
Db 4520 ATTTTCAATTCTGCCAGAGAAATATTCAAAATATATGACAGTGCCTGATAGAGATG 4579  
QY 1886 ggccaactaaatcagactcctcgtgtgcctgacccgtatctactagttcgtaatatgg 1945  
Db 4580 GGCCAACTAAATCAGATCCCTCGGCTGACCCCTATTACTAGTTCCGTAAATATGG 4639  
QY 1946 aagaagcttagcttcaaggactatggccctccctcaatcgtgttaaagaactctag 2005  
Db 4640 AGAAGATCTAGCTTCAGGACTCATGTGCCCTCTCTCATCTGCTACAAAGAACTGTAG 4699  
QY 2006 atcaaaaggaaacagataatgcaagaacaagagaaatgcatcctgttctctatgtt 2065  
Db 4700 ATCAAGAGGAAACACAGATATATGTCAGACAAAGAGAAATGTCATCTGTATTTTG 4759  
QY 2066 atggaacccgaagctgtgtactcacaagagaataatacagccttctcccaatccagctg 2125  
Db 4760 ATGGAACCCGAAGCTGTACTCTCACAAGAAATATCAACGTTTCTCTCCCAATCCAGTGG 4819  
QY 2126 gaggcagcttgaggatcccaagatcccaagctccaacatcatgacagcaatcaatgct 2185  
Db 4820 GAGTGCAGCTTGAGGATCCAGAGTTCCAAAGCTCCAAATATACAGCAATCATGGCT 4879  
QY 2186 atgttttgatagtttgcagttgcagttgttgcataagagtgagcaactgtaactc 2245  
Db 4880 ATGTTTGAATGTTTGCAGTTCAGTTCATTTGTTGCAATGAGGTGGCATACTGTAATTC 4939  
QY 2246 taagcatltgagacagacagtaacttccctctcgtctctctctctgataactcacaac 2305  
Db 4940 TAAACATTTGGAGCAACAGATGACTTCTTCTGTTCTTCTCTGATATACCTTTCAAC 4999  
QY 2306 acaaaatggtctatagaagacacactcaacctatccatctcaggggaacatcttca 2365  
Db 5000 ACAAAATGGTCTATGAAGACACACTCACCTATTCCTCAATTCAGGAGAAATCTCTTCA 5059  
QY 2366 tgtcagatgaaaaacccaggtctatgatactctgggtgtgccaacatcagacttgaga 2425  
Db 5060 TGTGATGAGAAACCCAGGTCTATGATTTCTGGGGTGCACAACTCAGACTTTGGAACA 5119  
QY 2426 gaggcatgacgccttaactgaaggttctctagttgtgacaagaacaactggtgataactag 2485  
Db 5120 GAGCATGACCGCTTACTGTAAGGTTCTAGTTTGTGACAAACACATGCGATATTATACG 5179  
QY 2486 aggcagcttatgaagatattccagcatactgctgagtaaaacaatgccaatgaaacaa 2545  
Db 5180 AGGACAGTTATGAAGATATTTTACGATATCTTGTGAGTAAAGAAATGCCATTGAACCAA 5239  
QY 2546 gaagcttctccagaatccaagacacctagcactagcacaagaacttaatgccaacc 2605

Db 5240 GAAGCTTCTCCAGAAATTCAAGACACCCCTAGCACTAGGCAAAAGCAATTTAATGCCACC 5299  
QY 2606 caccagctctgaaaacgcacatcaacaggaataactcgtactactcttcagtcagatacag 2665  
Db 5300 CACGAGTCTTGAAGCGCATACCGGAAATAACTCGTACTACCTCTCAGTCAATCATCAG 5359  
QY 2666 aggaatctgactatgatatccatcaactgtaagtgaagaagaagatttacaatt 2725  
Db 5360 AGGAATTTGACTATGATGATATCCATATCAGTTGAAATGAAAGAAAGATTTTACATTT 5419  
QY 2726 atgataagatgaaaatcagaagcccccgagcttccaagaagaacaacagacataatta 2785  
Db 5420 ATGATGAGAGATGAATAATCAGAGCCCCGACCTTTCAAAAAGAAACGACACTATTTTA 5479  
QY 2786 ttgctgcagtggaagagctcctgggattatgagatgagtagtctcccatatgttctaagaa 2845  
Db 5480 TTTGCTGAGTGAAGAGGCTCGGATTTATGGAGTAGAGTCCCCACATCTTTTAAGAA 5539  
QY 2846 acaaggtcagagtgagcagtgctccctcagttcaagaagaagtgtttccaagaattactg 2905  
Db 5540 ACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTTCCAGCAATTTACTG 5599  
QY 2906 atggctccttactcagcccttataccgttggaagacataatgaaacatttggactcctg 2965  
Db 5600 ATGGCTCCTTACTCAGCCCTTATACCGTGAAGAACTAAATGAAACATTGGGACTCTGG 5659  
QY 2966 ggccatatataagacgaagatgaaagataatacaatgtaactttcaagaatacagcct 3025  
Db 5660 GGCCATATATGAAGCAGAACTTAAGATATATCATGTGTAATCTTCAAGAAATCAGGCT 5719  
QY 3026 ctggtccctatctctcctcactcagcttacttcttaagaagaagatcagaagcagaag 3085  
Db 5720 CTGCTCCTTATTCCTTCTTATCTAGCTTATTTCTTATGAGGAAGATCAGAGCAAGAG 5779  
QY 3086 cagaacctagaaaaaacttgttcaagcctaaatgaacccaaacttacttltgaaagtgc 3145  
Db 5780 CAGAACCTAGAAAAAAGCTTTGTCAAGCCTAATGAACCAAACTTACTTTGAAAGTGC 5839  
QY 3146 aacatcataatggcaccactaaagatgatttgcactgcaaaagccttggctattctctg 3205  
Db 5840 AACATCATATGGCACCCCACTAAAGATGAGTTGACTGCAAGGCTTGCGCTTATTTCTGTG 5899  
QY 3206 atgttgacctgaaaagaatgtgcactcagccggaattggaccccttctgtctgcaca 3265  
Db 5900 ATGTTGACCTGGAAAAAGATGTGCATCAGGCTGATTTGACCCCTTGTGCTGCCACA 5959  
QY 3266 ctaacacactgaaacctgtctcaatggagacaagtgcacagtaacagaaattgtcctgttt 3325  
Db 5960 CTAACACACTGAACCTGCTCATGTGGAGACAAGTGACAGTACAGGAATTTGCTCGTTTY 6019  
QY 3326 tcacacatcttgaatgagacccaagaagctgtgtacttccactggaataatggaagaactgca 3385  
Db 6020 TCACCATCTTTGATGAGACCAAAAGCTGTACTTACATGAAATATGGAAGAAACTGCA 6079  
QY 3386 gggtccctcgtcaatcacaatggaatgccacttttaagagaatatgcttccatg 3445  
Db 6080 GGGCTCCCTGCAATATTCAGATGGAAGATGCCACTTTTAAAGAAATATATGCTTCCATG 6139  
QY 3446 caatcaatggtctacataatgatacactacctgtgtagtaatggtctcaagatcaaaaaga 3505  
Db 6140 CAATCAATGGCTTACATATATGATATACACTACCTGCTTATGTAATGGCTCAGAGTCAAAAGA 6199  
QY 3506 ttgagtgatctgtctcagacatgggcagcaatggaataatccatcttcatattcaattcagtg 3565  
Db 6200 TTTGATGGTATCTCTCAGCTGAGCTGGCAAGCAATGGAACATCTATTTCAATTTCAAGTG 6259  
QY 3566 gacatgtgtcactgtaacgaaaaaagaggtataaaatgacctgtacatactcactca 3625  
Db 6260 GACATGTGTTTACCTGTAAGAAAAAGAGAGTATTAATAATGCACTGTACAAATCTATATC 6319  
QY 3626 caagtgattttgagacagtggaatgtlthacaatccaaagcttgaatttggcgggtggaat 3685  
Db 6320 CAGGTGTTTGAAGACAGTGAATGTATACATCAAAAGCTGAATTTGGCGGTGGAAT 6379



|||||  
Db 444 AGGGAAGAAAGATGATGATAAGTCTTCCCTGGTGAAGCATTATGCTGGAGGTC 503  
Qy 748 ctgaagaagaatggtccaatggtccctctgaccacgtgcttactactatcatcttct 807  
Db 504 CTGAAGAAGAAATGGTCCATGACCTCTTGACCCACAGTGCCTTACTCTCATCTCTTCT 563  
Qy 808 catgtgagccctgtaaaagacttgatccagccctcatcttgagccctctagtatgata 867  
Db 564 CATGTGAGCCTGGTAAAGACTTGATTCAGGCCCTCATTTGAGGCCCTACTGTATGTGA 623  
Qy 868 gaaggagctctgcccgaagaaagacacagacactctgcaaaattatactacttctgct 927  
Db 624 GAAGGAGCTCTGGCCAAAGAAAGACACAGACCTTGACAAATTTATCTACTTTTCT 683  
Qy 928 gtatttgatgaaggaaagtctggaccctcagaacaagaactccctgatagcagatagg 987  
Db 664 GTATTGTGATGAAGGAAAAGTTGGCACTCAGAAACAAAGAACTCCTTGATGCAGATAGG 743  
Qy 988 gatctgcatctgctcgggcttggtccctaaatgcaacagctcaatggttatgtaacagg 1047  
Db 744 GATGCTGATCTGCTCGGGCTCGGCTTAATTCACACAGTCAATGTTATGTAAACAGG 803  
Qy 1048 tctctgcaaggctctgatatgatagcacagaaatcaatctctgcatgtagtgaatg 1107  
Db 804 TCTGTGCCAGGCTGTGATGTGATGCAAGCAAGAAATCAGTCTATTGGCATGTGATGGAAAG 863  
Qy 1108 ggcacacactcctgaagtgacatcaatctcctcgaagtgacacactctctgtgaagaa 1167  
Db 864 GGCACACTCTCTGAAGTGACACTCAATATCTCGAAGGTCCACACATTTCTTGAGGAAAC 923  
Qy 1168 catgcgcaaggctccttggaatctcgccaaataacttccctacttgctcaaaactctg 1227  
Db 924 CATGCCAGGCGCTCTTGGAATCTCCGCAATTACTTTCTTACTGCTCAACACTCTTG 983  
Qy 1228 atggaccttgagacagttctactglttctgtaatactcttccacacacatgtagatg 1287  
Db 984 ATGGACCTTGGACAGTTTCTACTGTTTGTGATATCTCTCCACCAACATGATGGCATG 1043  
Qy 1288 gaagacttatgcgaagtatgacagctgtccagagaagaccccaactacgaatgaataat 1347  
Db 1044 GAACCTTATGTCAAAAGTAAAGAGCTGTCCAGAGAACCCCAACTACAGATGAAAAATAT 1103  
Qy 1348 gaagaagcggaagactatgacgactctactctgacttgaaatgtagtgaagttt 1407  
Db 1104 GAAGAAGCGGAAGACTATGATGATGATCTTACTATTTCTGAAATGATGTGTAGGTTT 1163  
Qy 1408 gatgatgaacactcctctctcttaccataatctgctcaagtctgcaagaagacactaa 1467  
Db 1164 GATGATGACACTCTCTCTCTTATCCAAATTCGCTCAGTTGCCAAGAAGCATCTTAA 1223  
Qy 1468 acttggtgacttatactatgctgctggaagaaggaactggtgactatgctccttactc 1527  
Db 1224 ACTTGGGTACATTTACATTTGCTGGAAGAGAGACTGGGACTATGCTCTCTTGTCTCTC 1283  
Qy 1528 gcccgcgatgcagaagttataaagaatcaatttgaaacaatgcccctcagcgaatcgt 1587  
Db 1284 GCCCCGATGACAAAGTTATTAAGATCAATATTGTAACAATGCGCCCTCAGCGGATGGT 1343  
Qy 1588 aggaagtaacaaaagtcgacttatactgcatcacacagatgaaaccttaagaacctgaa 1647  
Db 1344 AGGAAGTAACAAAAAGTCCGATTTATGGCATACACAGATGAACCTTTAAGACTCGTAA 1403  
Qy 1648 gctattcaagcatgaatacagaactctgggaacttacttacttattggtgaagtgtgaaca 1707  
Db 1404 GCTATTGTAGCTGATGATGAGATCTTGAGACTTTACTTTATGCGGAAGTTGGAGACAA 1463  
Qy 1708 ctgttgattatatttaagaatcaagcaagcagacacataacatctaccctcagaagatc 1767  
Db 1464 CTGTGTGATTAATTTAAAGATCAAGCAAGACCATTAATCAATCTACCTCAGCGAATC 1533  
Qy 1768 actgattcgcgtccttctgattacaaagatctacaaaggtgtaaacacttgaagat 1827  
|||||

Db 1524 ACTGATGTCCCTCTTGTATTCAAGGAGATTACCAAAAAGTGTAAACATTTCAGGAT 1583  
Qy 1828 ttccaattctgcagagagaataattcaataataatggaagtgactgtgaagaatggg 1887  
Db 1584 TTTCCAAATTTGCCAGGAGGAATATTCAAATATTAATGACAGTGTGAGAGATGGG 1643  
Qy 1888 ccaactaaatcagatccctcgggtgcctgacccgcatctactctagtcttctgtaataatg 1947  
Db 1644 CCAACTAATATAGATCCTCGGTGCTGACCCGCTATTACTTACTTTGCTTAATATGGAG 1703  
Qy 1948 agagatcctagcttaaggactcatctggccctctcctcatctgctcaaaagaatctgtagat 2007  
Db 1704 AGAGATCTAGCTTCAGAGCTCATTTGGCCCTCTCCATCTGCTCAAAAGAAATCTGTAGAT 1763  
Qy 2008 caaagaagaaacagataatgtaacagaagaagtcatcctglttctgtaattggtat 2067  
Db 1764 CAAAGAGGAAAACGAAATATGTACAGAAAGAGATGCATCTGTTTCTGTATTGTAT 1823  
Qy 2068 gagaaccgaagctggtacactcacagagaataatacaacgcttctcccaatccagctgga 2127  
Db 1824 GAGAACCGAAGCTGTGATCTACAGAGATATACAGCTTTCTCCCAATTCAGCTGGA 1883  
Qy 2128 gtgcaagcttgaggaatccagagttccaaagctcacaacalcatgcaagaatcaatgtagat 2187  
Db 1884 GTGCAGCTTGAGATGCCAGATTCCAAGCTCCAAACATCATGACAGCATCAATGGCTAT 1943  
Qy 2188 gtctttagatgttgcagttgtcgaattgttgcatagaagtgtgcatactgtaattccta 2247  
Db 1944 GTTTTATATATTGTGACGTTGTGAGTTGTTGATGATGAGGAGCATCTGTGATCAATCTA 2003  
Qy 2248 agcatggaagacagactggaactccttctgctctctctctctgataacttcaaaacac 2307  
Db 2004 AGCATTTGGAGACAGACTGACTTCTTCTTCTTCTTCTTCTGATATACCTTTAAACAC 2063  
Qy 2308 aaaaatgltctaagaacacactacactatcccatctcagaagaagaactgltctcatg 2367  
Db 2064 AAATGTGCTATGAAGACACACTCACCTATTTCCATTTCCAGAGAAACGTCTTCATG 2123  
Qy 2368 tctgtggaataacccaagttctatggaattctgggtgtgccaaactaagactctcgaacaaga 2427  
Db 2124 TCGATGGAATAACCCAGGTCTATGATTTCTGGGGTCCACAACTCAGACTTTCGGAACAGA 2183  
Qy 2428 ggcataccgcctctactgaaagttctcagttgtgtgaagaagaactggtgattatctagag 2487  
Db 2184 GGCATGACCCGCTTACTGAAAGTTTCTTAGTTGTACAAAGAACCTGTGATTTATTCAG 2243  
Qy 2488 gacagttatgaagataattccaagcaactctgctgataaacaatgcatctgaaccaaga 2547  
Db 2244 GACAGTTATGAAGATATTTTCAGCATCTGCTGATGATTAACATGTCATTTGAACCAAGA 2303  
Qy 2548 agcttctcccaagaattcaagaacacctagaactagcaaaagaacttaatgccaacca 2607  
Db 2304 AGCTTCTCCCAAGATTCAAGCACCTTAGCACTGAGCAAAAAGCAATTAATGCGCACCCA 2363  
Qy 2608 ccaagttctgaagaagcacaatcagggaaataactgtaactactctctcagtaagatcaagag 2667  
Db 2364 CCACTCTTGAACCCCACTCAAGCGGAATTAATCTGCTACTCTTTCAGTCAATCAAGAG 2423  
Qy 2668 gaaattgactatgataatacatatcagttgaaatgaaagaagaagaatttgacattat 2727  
Db 2424 GAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2483  
Qy 2728 gatgaagatgaataatcagagcccccgagcttcaaaagaagaacagacataattatc 2787  
Db 2484 GATGAGATGAAATTCAGAGCCCCCGAGCTTCAAAAAGAAACACACACTATTTTATTT 2543  
Qy 2788 gctgcaatggaagagccttggaattatggaatgtagtgcacacacatggttctcaagaagac 2847  
Db 2544 GCTGCAATGAGAGGCTCTGGGATTTATGGATGAGTGAAGTCTCCACATGTTCTTAAGAAC 2603  
Qy 2848 agggctcagagttggcagttgctcctcagttcaagaagaattgttttccaggaattacagat 2907  
Db 2604 AGGGCTCAGAGTGGCAGTGTCCCTCACTTCAAGAAAGTTGTTTCCAGGAATTTACTGAT 2663  
|||||

2908 gctcccttaacagcccttaacgltgagaactaaatgaacatttggacccctggg 2967  
|||||  
2664 ggctcccttaacagcccttaacgltgagaactaaatgaacatttggacccctggg 2723  
|||||  
2968 ccatataaagaagaagttgaagaataatcatgttaaccttcaagaatcaggccctc 3027  
|||||  
2724 ccatatattaaagcagaagttgaagaataatcatgttaaccttcaagaatcaggccctc 2783  
|||||  
3028 gctcccttaacagcccttaacgltgagaactaaatgaacatttggacccctggg 3087  
|||||  
2784 ggctcccttaacagcccttaacgltgagaactaaatgaacatttggacccctggg 2843  
|||||  
3088 gaacattgaagaagaagttgaagaataatcatgttaaccttcaagaatcaggccctc 3147  
|||||  
2844 gaacattgaagaagaagttgaagaataatcatgttaaccttcaagaatcaggccctc 2903  
|||||  
3148 catcatatgagcagcccaataagaatgttgaactgcaagccctgggttattctctgat 3207  
|||||  
2904 catcatatgagcagcccaataagaatgttgaactgcaagccctgggttattctctgat 2963  
|||||  
3208 gttgacccctgagaagaagttgacacagccctgattgagccctctgtctgcaagcc 3267  
|||||  
2964 gttgacccctgagaagaagttgacacagccctgattgagccctctgtctgcaagcc 3023  
|||||  
3268 aacacacatgaacccctgctcagagagagacagtgacagagaaatttgcctgtcttcc 3327  
|||||  
3024 aacacacatgaacccctgctcagagagagacagtgacagagaaatttgcctgtcttcc 3083  
|||||  
3328 aacacacatgaacccctgctcagagagagacagtgacagagaaatttgcctgtcttcc 3387  
|||||  
3084 aacacacatgaacccctgctcagagagagacagtgacagagaaatttgcctgtcttcc 3143  
|||||  
3388 gttgacccctgagaagaagttgacacagccctgattgagccctctgtctgcaagcc 3447  
|||||  
3144 ggctcccttaacagcccttaacgltgagaactaaatgaacatttggacccctggg 3203  
|||||  
3448 atcaatgtgctacataatgatacactacagctgctgaatgaatgctcagagatcaaaagatt 3507  
|||||  
3204 atcaatgtgctacataatgatacactacagctgctgaatgaatgctcagagatcaaaagatt 3263  
|||||  
3508 cgaatgtatctgctcag 3567  
|||||  
3264 cgaatgtatctgctcag 3323  
|||||  
3568 catgtatctgctcag 3627  
|||||  
3324 catgtatctgctcag 3383  
|||||  
3628 ggtgttttggag 3687  
|||||  
3384 ggtgttttggag 3443  
|||||  
3688 ctatattggag 3747  
|||||  
3444 ctatattggag 3503  
|||||  
3748 tgcagagctccctggag 3807  
|||||  
3504 tgcagagctccctggag 3563  
|||||  
3808 ggaacataatgag 3867  
|||||  
3564 ggaacataatgag 3623  
|||||  
3868 gctctgag 3927  
|||||  
3624 gctctgag 3683  
|||||  
3928 attcag 3987  
|||||  
3684 attcag 3743  
|||||

3988 ttatcatcatgataagcttctgataagagagagagagagagagagagagagagagagagagag 4047  
|||||  
3744 ttatcatcatgataagcttctgataagagagagagagagagagagagagagagagagagagag 3803  
|||||  
4048 ggaacataatgag 4107  
|||||  
3804 ggaacataatgag 3863  
|||||  
4108 aacccctcaatattgctcagataacccgttggaccccaatcatatgaatctgagc 4167  
|||||  
3864 aacccctcaatattgctcagataacccgttggaccccaatcatatgaatctgagc 3923  
|||||  
4168 actcttcagatgag 4227  
|||||  
3924 actcttcagatgag 3983  
|||||  
4228 gagagtaagaagaataatcagatgagagagagagagagagagagagagagagagagagagag 4287  
|||||  
3984 gagagtaagaagaataatcagatgagagagagagagagagagagagagagagagagagagag 4043  
|||||  
4288 gccacctgctcctctcaaaagctcagactcactcacaagagagagagagagagagagagagagag 4347  
|||||  
4044 gccacctgctcctctcaaaagctcagactcactcacaagagagagagagagagagagagagagag 4103  
|||||  
4348 cctcagatgag 4407  
|||||  
4104 cctcagatgag 4163  
|||||  
4408 aacag 4467  
|||||  
4164 acag 4223  
|||||  
4468 ctcaatctcagaag 4527  
|||||  
4224 ctcaatctcagaag 4283  
|||||  
4528 aaggttttcaag 4587  
|||||  
4284 aaggttttcaag 4343  
|||||  
4588 ttactgactgctcactcag 4647  
|||||  
4344 ttactgactgctcactcag 4403  
|||||  
4648 atgag 4699  
|||||  
4404 atgag 4455  
|||||  
RESULT 4  
US-09-470-618-14  
: Sequence 14, Application US/09470618  
: Patent No. 6200560  
: GENERAL INFORMATION:  
: APPLICANT: Coulo, Linda B.  
: APPLICANT: Colosi, Peter C.  
: TITLE OF INVENTION: Adeno-Associated Vectors for Expression of Factor VIII  
: TITLE OF INVENTION: By Target Cells  
: FILE REFERENCE: Avigen-04082  
: CURRENT APPLICATION NUMBER: US/09/470,618  
: EARLIER FILING DATE: 1999-12-22  
: EARLIER APPLICATION NUMBER: 09/364,862  
: EARLIER FILING DATE: 1999-07-30  
: EARLIER APPLICATION NUMBER: 60/125,974  
: EARLIER FILING DATE: 1999-03-24  
: EARLIER APPLICATION NUMBER: 60/104,994  
: EARLIER FILING DATE: 1998-10-20  
: NUMBER OF SEQ ID NOS: 15  
: SOFTWARE: PatentIn Ver. 2.0  
: SEQ ID NO 14  
: LENGTH: 4999  
: TYPE: DNA  
: ORGANISM: Artificial Sequence

```
; FEATURE:
;
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-470-618-14
```

|                       |   |              |              |             |
|-----------------------|---|--------------|--------------|-------------|
| Query Match           | 90.9%   | Score 4330.8 | DB 4         | Length 4599 |
| Best Local Similarity | 99.0%   | Prod. No. 0  |              |             |
| Matches 4384          | Conservative  | 0            | Mismatches 2 | Indels 42   |
|                       |   |              |              | Gaps 1      |
| QY 266                | ccacatgccaataagaagctctccacctgctcttcctctgtgctcttttgagatcgtcctta        | 325          |              |             |
| Db 411                | ccccatgccaataagaagctctccacctgctcttcctctgtgctcttttgagatcgtcctta        | 470          |              |             |
| QY 326                | gtgcacacagaagaatactactcctgggtgcagtcggaactgcatactggaacatactgaacg       | 385          |              |             |
| Db 471                | gtgcacacagaagaatactactcctgggtgcagtcggaactgcatactggaacatactgaacg       | 530          |              |             |
| QY 386                | atctctgtgagatcgtcctctgtgcagcaaatctctctccagagtcaccaatctttccat          | 445          |              |             |
| Db 531                | atctctgtgagatcgtcctctgtgcagcaaatctctctccagagtcaccaatctttccat          | 590          |              |             |
| QY 446                | tcaaacacttcagctgtgacacaaaagactcgtttgtgaattcaacgggtcaacctttca          | 505          |              |             |
| Db 551                | tcaaacacttcagctgtgacacaaaagactcgtttgtgaattcaacgggtcaacctttca          | 650          |              |             |
| QY 506                | acatcgtcctaagccaagcgccaccttgcagtcgggtctgtcagtcctacccaaccagtcgag       | 565          |              |             |
| Db 651                | acatcgtcctaagccaagcgccaccttgcagtcgggtctgtcagtcctacccaaccagtcgag       | 710          |              |             |
| QY 566                | tttatgatacagctgtgtcatctacaacttaagaacatggtctccacatccctgtcagcttcat      | 625          |              |             |
| Db 711                | tttatgatacagctgtgtcatctacaacttaagaacatggtctccacatccctgtcagcttcat      | 770          |              |             |
| QY 626                | ctgtgtgagctatcccaactcggaagaagctctctggggagcgcggaatgatatgatacagcagtc    | 685          |              |             |
| Db 771                | ctgtgtgagctatcccaactcggaagaagctctctggggagcgcggaatgatatgatacagcagtc    | 830          |              |             |
| QY 686                | aaagggagaaagaagatgataaagctcttcctcctgtgtggaagccaatacatatgctgcgcaag     | 745          |              |             |
| Db 831                | aaagggagaaagaagatgataaagctcttcctcctgtgtggaagccaatacatatgctgcgcaag     | 890          |              |             |
| QY 746                | tccctgaagaagaatggtgtccaatgagctctctggcccacacttgctcttaaccattacatcttt    | 805          |              |             |
| Db 891                | tccctgaagaagaatggtgtccaatgagctctctggcccacacttgctcttaaccattacatcttt    | 950          |              |             |
| QY 806                | ctctatgtgagacctgtgtaaaaagacttgaaattcgaagcctcatctggagccctactagatgta    | 865          |              |             |
| Db 951                | ctctatgtgagacctgtgtaaaaagacttgaaattcgaagcctcatctggagccctactagatgta    | 1010         |              |             |
| QY 866                | gagagaaggagctctgtgccaaggaaaaagacacagacactgtgcacaaaatttactactctttg     | 925          |              |             |
| Db 1011               | gagagaaggagctctgtgccaaggaaaaagacacagacactgtgcacaaaatttactactctttg     | 1070         |              |             |
| QY 926                | ctgtattttgataagaaagggaaaaagtttgagcctcgaacaaacaaagaaaccccttgatcgaggta  | 985          |              |             |
| Db 1071               | ctgtattttgataagaaagggaaaaagtttgagcctcgaacaaacaaagaaaccccttgatcgaggta  | 1130         |              |             |
| QY 986                | gggaatgctgcatctcgtctcgggcctctgcccataaatgacacagctcaatggttatgtaaca      | 1045         |              |             |
| Db 1131               | gggaatgctgcatctcgtctcgggcctcctgagcctaataatgcacacagctcaatggttatgtaaca  | 1190         |              |             |
| QY 1046               | ggctctctgcacaggtcctgatttgatgagccacagaagaatcaagctctattggcgcatgtgattgaa | 1105         |              |             |
| Db 1191               | ggctctctgcacaggtcctgatttgatgagccacagaagaatcaagctctattggcgcatgtgattgaa | 1250         |              |             |
| QY 1106               | tgggcacaacaccccggaatgagcactcaaatctccctcgaaagtcacaacacattctctggagga    | 1165         |              |             |
| Db 1251               | tgggcacaacaccccggaatgagcactcaaatctccctcgaaagtcacaacacattctctggagga    | 1310         |              |             |
| QY 1166               | acacatgcacagcgctcctgtgaaatctgcgccaataactttcccttaactgtccaacaactct      | 1225         |              |             |
| Db 1311               | acacatgcacagcgctcctgtgaaatctgcgccaataactttcccttaactgtccaacaactct      | 1370         |              |             |

|    |      |  |      |
|----|------|--|------|
| QY | 1226 | tgaagaccttggacagttctctacggttttgltacatactctccaccaacatgatgta         | 1285 |
| Dp | 1371 | tgaatgacaccttgagcagttctctactagttttgtgcatactctccaccaacatgatgta      | 1430 |
| QY | 1286 | tggaagacttatgtcaaaagttagaagcgttcctcagaagaaacccaactacgaatgaanaata   | 1345 |
| Dp | 1431 | tggaagacttatgtcaaaagttagaagcgttcctcagaagaaacccaactacgaatgaanaata   | 1490 |
| QY | 1346 | atgaaagaagcggaaagactaatgatgatgtatctctgcatctcgaatgtgtgcagat         | 1405 |
| Dp | 1491 | atgaaagaagcggaaagactaatgatgatgtatctctgcatctcgaatgtgtgtgcagat       | 1550 |
| QY | 1406 | tgtatgtagacaaactctctccttccattccaatctcgctcagttgccaagaagcatccta      | 1465 |
| Dp | 1551 | tgtatgtagacaaactctctccttccattccaatctcgctcagttgccaagaagcatccta      | 1610 |
| QY | 1466 | aaacttgggtatcatatcatattgctgcctggaagagaagaaactgggaactatgtcccttagtcc | 1525 |
| Dp | 1611 | aaacttgggtatcatatcatattgctgcctggaagagaagaaactgggaactatgtcccttagtcc | 1670 |
| QY | 1526 | tcgccccgagatgaagaagatctataaaagtcaaatctttgacaatgtgcccctcagcgatg     | 1585 |
| Dp | 1671 | tcgccccgagatgaagaagatctataaaagtcaaatctttgacaatgtgcccctcagcgatg     | 1730 |
| QY | 1586 | gtagaagaatcacaaaaagctccgaattltaggcatacacagaatgaaaccttlaagactctg    | 1645 |
| Dp | 1721 | gtagaagaatcacaaaaagctccgaatttaggcatacacagaatgaaaccttlaagactctg     | 1790 |
| QY | 1646 | aagctatcacagcatgaaatcagaactcttggacaccttactctatgaggggaagtgtgagaca   | 1705 |
| Dp | 1791 | aagctatcacagcatgaaatcagaactcttggacaccttactctatgaggggaagtgtgagaca   | 1850 |
| QY | 1706 | cacgttggattatattttaaaatccaagaagcagagaccataaactctaacctccacgaa       | 1765 |
| Dp | 1851 | cacgttggattatattttaaaatccaagaagcagagaccataaactctaacctccacgaa       | 1910 |
| QY | 1766 | tcactgtatgtccgctctcttctatccaaggagattaccaaaaggttlaaaactcttgaag      | 1825 |
| Dp | 1911 | tcactgtatgtccgctctcttctatccaaggagattaccaaaaggttlaaaactcttgaag      | 1970 |
| QY | 1826 | attttccaatctgccaaggagaataattccaataataatggaagtgactgtagaagatg        | 1885 |
| Dp | 1971 | attttccaatctgccaaggagaataattccaataataatggaagtgactgtagaagatg        | 2030 |
| QY | 1886 | ggccaactaaatcagaatccctccggtgctctgacccgccttaccctagtttcgttaataag     | 1945 |
| Dp | 2031 | ggccaactaaatcagaatccctccggtgctctgacccgccttaccctagtttcgttaataag     | 2090 |
| QY | 1946 | agaagaatctagatctcagaagcatattggccctctctcactcgtcatacaaaagatctgag     | 2005 |
| Dp | 2091 | agaagaatctagatctcagaagcatattggccctctctcactcgtcatacaaaagatctgag     | 2150 |
| QY | 2006 | atcaaaagagaaacagataatgtctgacacaagaaggaaigtcatccgttttctgtatttg      | 2065 |
| Dp | 2151 | atcaaaagagaaacagataatgtctgacacaagaaggaaigtcatccgttttctgtatttg      | 2210 |
| QY | 2066 | atggaagaccgaagctggttactccaagagaataataaaagctttctcccaatccagctg       | 2125 |
| Dp | 2211 | atggaagaccgaagctggttactccaagagaataataaaagctttctcccaatccagctg       | 2270 |
| QY | 2126 | gagcgaagctgtgaagatccagagttccaaagccccaacaatactgacaagatcaatggct      | 2185 |
| Dp | 2271 | gagcgaagctgtgaagatccagagttccaaagccccaacaatactgacaagatcaatggct      | 2330 |
| QY | 2186 | atgtttttgatagtttgcaagtgtgcaagtgtgtttgcatgaggtggtacatactgtacatc     | 2245 |
| Dp | 2331 | atgtttttgatagtttgcaagtgtgcaagtgtgtttgcatgaggtggtacatactgtacatc     | 2390 |
| QY | 2246 | taagcatctgagaacacagacactgacttctcttcgtctcttctctctgataataactccaac    | 2305 |
| Dp | 2391 | taagcatctgagaacacagacactgacttctcttcgtctcttctctctgataataactccaac    | 2450 |
| QY | 2306 | acaaaatggtcgtatgaaagacacactcaacctatcccatcttccaagataaactgtcttca     | 2365 |

|||||  
Db 2451 acaaaatgctatgaaacacacactcacccctatcccaattctcagagaaatgtcttca 2510  
QY 2366 tctcgatgaaacccagctcatgatctctgggtgcccacactcagaactctcgaaaca 2425  
Db 2511 tgtcgatgaaacccagctcatgatctctgggtgcccacactcagaactctcgaaaca 2570  
QY 2426 gaggcatgacccgcttacttgaaaggttctagtgtgcaacaagaacatggtatattacg 2485  
Db 2571 gaggcatgacccgcttacttgaaaggttctagtgtgcaacaagaacatggtatattacg 2630  
QY 2486 aggcaggtatagaagatatcttcagcatactctgtgagttaaaaaaatgcatcttgaaacca 2545  
Db 2631 aggcaggtatagaagatatcttcagcatactctgtgagttaaaaaaatgcatcttgaaacca 2690  
QY 2546 gaagcttctcccaagattccaagacacccctagcactagcaaaagaatttaatgcccac 2605  
Db 2691 gaagcttctcccaagat-----c 2708  
QY 2606 caccagttctgaaacgcatcaacggaataaactcgtactactcttcagtcagaatcaag 2665  
Db 2709 caccagttctgaaacgcatcaacggaataaactcgtactactcttcagtcagaatcaag 2768  
QY 2666 aggaatctgactatgatatataccataatcagttgaaatgaagaagaagatttgacatt 2725  
Db 2769 aggaatctgactatgatatataccataatcagttgaaatgaagaagaagatttgacatt 2828  
QY 2726 atgattgaagatgaataatcagaagcccccagcttccaagaagaacacacacacttcta 2785  
Db 2829 atgattgagatgaataatcagaagcccccagcttccaagaagaacacacacacttcta 2888  
QY 2786 tctgctgcaatgagagagctctgagataatggagatgagctcccaactgttctaaga 2845  
Db 2889 tctgctgcaatgagagagctctgagataatggagatgagctcccaactgttctaaga 2948  
QY 2846 acaaggtctcagaatgagcagtgctccctcagttcaagaagaatgtttccaaggaatttaccg 2905  
Db 2949 acaaggtctcagaatgagcagtgctccctcagttcaagaagaatgtttccaaggaatttaccg 3008  
QY 2906 atggtctcttaccacagccctataacgctgagaaactaaatgaacatttggaaactctcg 2965  
Db 3009 atggtctcttaccacagccctataacgctgagaaactaaatgaacatttggaaactctcg 3068  
QY 2966 ggcacataataagacagaagctggaagataatataatcagtgtaacttccagaaatcagacct 3025  
Db 3069 ggcacataataagacagaagctggaagataatataatcagtgtaacttccagaaatcagacct 3128  
QY 3026 ctgctccctattctcttctatcttagccttatcttctatgagaaatcagaagcagaag 3085  
Db 3129 ctgctccctattctcttctatcttagccttatcttctatgagaaatcagaagcagaag 3188  
QY 3086 cagaacctaagaaaaactctgccaagcctaaatgaaccaaacttacttctgaaagtgc 3145  
Db 3189 cagaacctaagaaaaactctgccaagcctaaatgaaccaaacttacttctgaaagtgc 3248  
QY 3146 aacatacataagcaccacataaagatgagcttgactgcaaaagcctgggcttaattctctg 3205  
Db 3249 aacatacataagcaccacataaagatgagcttgactgcaaaagcctgggcttaattctctg 3308  
QY 3206 atgtctgacctgaaaaagatgtgcaactcagagcctgattggacccttctgtgtccaca 3265  
Db 3309 atgtctgacctgaaaaagatgtgcaactcagagcctgattggacccttctgtgtccaca 3368  
QY 3266 ctacacactgaacccctctcatgtgagacaaagtacacagttacagaatctgtctgttt 3325  
Db 3369 ctacacactgaacccctctcatgtgagacaaagtacacagttacagaatctgtctgttt 3428  
QY 3326 tcaacatcttctgataagaccaaagctgtgtactctcactgtaaaataatggaaagaaactgca 3385  
Db 3429 tcaacatcttctgataagaccaaagctgtgtactctcactgtaaaataatggaaagaaactgca 3488  
QY 3386 gggctccctgcaataatcagatggaagatcccaatttaagaagaattatcgcttccatg 3445  
|||||

Db 3489 gggctccctgcaataatcagatggaagatcccaatttaagaagaattatcgcttccatg 3548  
QY 3446 caatcaatgctacataaataatgatacactacactgcttgaataatgctcagagatcaaaaga 3505  
Db 3549 caatcaatgctacataaataatgatacactacactgcttgaataatgctcagagatcaaaaga 3608  
QY 3506 ttcgattgatactgctcagacatgggacgaataatgaataatccatcttaattcaatg 3565  
Db 3609 ttcgattgatactgctcagacatgggacgaataatgaataatccatcttaattcaatg 3668  
QY 3566 gacatgtgttcactgtacgaaaaaagagagatataaataatgacactgtacatcttcac 3625  
Db 3669 gacatgtgttcactgtacgaaaaaagagagatataaataatgacactgtacatcttcac 3728  
QY 3626 caggtgtttttagaacagtgaataatgttaccataccaaagcggagatttgcgggttggaat 3685  
Db 3729 caggtgtttttagaacagtgaataatgttaccataccaaagcggagatttgcgggttggaat 3788  
QY 3686 gcttaattgagcagcatcacaatgctggaatgagcacactttctcgtgtacagcaata 3745  
Db 3789 gcttaattgagcagcatcacaatgctggaatgagcacactttctcgtgtacagcaata 3848  
QY 3746 agtctcagaatccctctgggaatggttcttgagacattataggaatttccagattacagct 3805  
Db 3849 agtctcagaatccctctgggaatggttcttgagacattataggaatttccagattacagct 3908  
QY 3806 caggaacaatataagacagctgggcccacaaagctggccagactcattatccggatcaatca 3865  
Db 3909 caggaacaatataagacagctgggcccacaaagctggccagactcattatccggatcaatca 3968  
QY 3866 atgcttgagacacaaagagaccttctcttgatcaaaagtgtgactgttgcacaaatga 3925  
Db 3969 atgcttgagacacaaagagaccttctcttgatcaaaagtgtgactgttgcacaaatga 4028  
QY 3926 ttattcacggacataaagcccagggctgccgctcagaagtcttcagacctcacaatcttc 3985  
Db 4029 ttattcacggacataaagcccagggctgccgctcagaagtcttcagacctcacaatcttc 4088  
QY 3986 agttatcatcatgatatgctcttgatggagaagctggcagacttatcgaaggaattcca 4045  
Db 4089 agttatcatcatgatatgctcttgatggagaagctggcagacttatcgaaggaattcca 4148  
QY 4046 ctggaaccttaatggtcttctcttgccaatgttgatcaatctcggagataaacaataatct 4105  
Db 4149 ctggaaccttaatggtcttctcttgccaatgttgatcaatctcggagataaacaataatct 4208  
QY 4106 ttaacctccaatataatgctcgataacatccgcttggacccaactcattatagcatcttgc 4165  
Db 4209 ttaacctccaatataatgctcgataacatccgcttggacccaactcattatagcatcttgc 4268  
QY 4166 gcaactcttcagatgagtgtgattggctgtgattaaatagcttgacagatgcatctggaa 4225  
Db 4269 gcaactcttcagatgagtgtgattggctgtgattaaatagcttgacagatgcatctggaa 4328  
QY 4226 tggagagttaaaagcaatcagatgacagatctaactgtcttaacttaacaaatgct 4285  
Db 4329 tggagagttaaaagcaatcagatgacagatctaactgtcttaacttaacaaatgct 4388  
QY 4286 ttggcaactgtgtctctcacaagcttgacttcaactccaaggaggagatgactcctgga 4345  
Db 4389 ttggcaactgtgtctctcacaagcttgacttcaactccaaggaggagatgactcctgga 4448  
QY 4346 gacctcaggtgataatccaaaagatgtgctgcaagtgaattccagaagaacataatgaag 4405  
Db 4449 gacctcaggtgataatccaaaagatgtgctgcaagtgaattccagaagaacataatgaag 4508  
QY 4406 tcaagaggttaactcacaaggttaaaatctctgtcttaacagacatgtatgtgaagagt 4465  
Db 4509 tcaagaggttaactcacaaggttaaaatctctgtcttaacagacatgtatgtgaagagt 4568  
QY 4466 tctcatctccaagcagtaagaatgagcctaactgagactcttcttccaagatggcaag 4525  
Db 4569 tctcatctccaagcagtaagaatgagcctaactgagactcttcttccaagatggcaag 4628  
|||||



QY 4526 taaagttttcaagggaatcaaacactcttcacacctgtgtgtaaacctcttagaccac 4585  
|||||  
Db 4629 taaagttttcagggaatcaaacactcttcacacctgtgtgtaaacctcttagaccac 4688  
QY 4586 cgttactgactcgctactccttcgaattacaccccaagagtgtgtgacacagattgtccctga 4645  
|||||  
Db 4689 cgttactgactcgctactccttcgaattacaccccaagagtgtgtgacacagattgtccctga 4748  
QY 4646 ggtatgaggtgttctggtgtcgtgacagagactcttactctcagac 4693  
|||||  
Db 4749 ggtatgaggtgttctggtgtcgtgacagagactcttactctcagac 4796  
  
RESULT 5  
US-09-364-862-14  
; Sequence 14, Application US/09364862  
; Patent No. 6221349  
; GENERAL INFORMATION:  
; APPLICANT: Coloto, Linda B.  
; APPLICANT: Colosi, Peter C.  
; TITLE OF INVENTION: ADEMO-ASSOCIATED VECTORS FOR EXPRESSION OF FACTOR VIII  
; TITLE OF INVENTION: BY TARGET  
; FILE REFERENCE: AVIGEN-03743  
; CURRENT APPLICATION NUMBER: US/09/364,862  
; EARLIER APPLICATION NUMBER: 60/125,974  
; EARLIER FILING DATE: 1999-03-24  
; EARLIER APPLICATION NUMBER: 60/104,994  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 14  
; LENGTH: 4999  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-364-862-14  
  
Query Match 90.9%; Score 4330.8; DB 4; Length 4999;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 4384; Conservative 0; Mismatches 2; Indels 42; Gaps 1;  
  
QY 266 ccacatgcaaatagagctctccacctgtcttctctgtgcttcttgagattctgttta 325  
|||||  
Db 411 ccacatgcaaatagagctctccacctgtcttctctgtgcttcttgagattctgttta 470  
QY 326 gtgcacacgaagaatatactactgtgtgcagtggaactgttcatatggagctataatgcaagt 385  
|||||  
Db 471 gtgcacacgaagaatatactactgtgtgcagtggaactgttcatatggagctataatgcaagt 530  
QY 386 atctcggtgagcgtcgtgtgagcgaagattctctctctagagtgccaatacttctccat 445  
|||||  
Db 531 atctcggtgagcgtcgtgtgagcgaagattctctctctagagtgccaatacttctccat 520  
QY 446 tcaaacctcagtcgtgtacaaaagactctgtttgtagaatcaccaggtcaccttttca 505  
|||||  
Db 591 tcaaacctcagtcgtgtacaaaagactctgtttgtagaatcaccaggtcaccttttca 650  
QY 506 acatgcctaagcgaagccacccctgtagtgcctgtagtcttaccatccagagcgtgag 565  
|||||  
Db 651 acatgcctaagcgaagccacccctgtagtgcctgtagtcttaccatccagagcgtgag 710  
QY 566 ttatgatatacgtgtcatatcacattaaagacatggcttcccatctctgcaagcttcacg 625  
|||||  
Db 711 ttatgatatacgtgtcatatcacattaaagacatggcttcccatctctgcaagcttcacg 770  
QY 626 ctgtgtgtgtatctctactgtgaaagctctcgaggagactgaatatgatgcagaacagtc 685  
|||||  
Db 771 ctgtgtgtgtatctctactgtgaaagctctcgaggagactgaatatgatgcagaacagtc 830

QY 686 aagggagaaagaagatgataaagttcttccctgtgtggaagccatataatgtctgacag 745  
|||||  
Db 831 aagggagaaagaagatgataaagttcttccctgtgtggaagccatataatgtctgacag 800  
QY 746 tcttgaagaagaatgggtccaatggtccctctgacccactgtgtccttactactacatacttt 805  
|||||  
Db 891 tcttgaagaagaatgggtccaatggtccctctgacccactgtgtccttactactacatacttt 950  
QY 806 ctcatgttgacccgtgttaaaagacttgaattcagggcctcatgtgagccctactatgata 865  
|||||  
Db 951 ctcatgttgacccgtgttaaaagacttgaattcagggcctcatgtgagccctactatgata 1010  
QY 866 gagaagaagagctgtgcccagaagaaacacagagccttgacaaattattracttctttg 925  
|||||  
Db 1011 gagaagaagagctgtgcccagaagaaacacagagccttgacaaattattracttctttg 1070  
QY 926 ctgtatttgaagaagaagtttgacactcagaacaaagaactccttgatcagagata 985  
|||||  
Db 1071 ctgtatttgaagaagaagtttgacactcagaacaaagaactccttgatcagagata 1130  
QY 986 gggaatcgtcatctgtcgtggtcgtggtcctaataatgcacacagtcaatggtatgataa 1045  
|||||  
Db 1131 gggaatcgtcatctgtcgtggtcgtggtcctaataatgcacacagtcaatggtatgataa 1190  
QY 1046 ggtctctgcaggtctgattgtagtgcacaggaatcagctcttggcatgtattgga 1105  
|||||  
Db 1191 ggtctctgcaggtctgattgtagtgcacaggaatcagctcttggcatgtattgga 1250  
QY 1106 tgggacacacatccttgaagtgacactcaaatctctcgtgaaggttcacacatctctgttagga 1165  
|||||  
Db 1251 tgggacacacatccttgaagtgacactcaaatctctcgtgaaggttcacacatctctgttagga 1310  
QY 1166 accatgcgcagcgtctcctgtggaatctcgcgaataacttccctactgtcacaacactct 1225  
|||||  
Db 1311 accatgcgcagcgtctcctgtggaatctcgcgaataacttccctactgtcacaacactct 1370  
QY 1226 ttgatgaccttggacaggttctctactgtttgtgatctctctccccaacagatgggca 1285  
|||||  
Db 1371 ttgatgaccttggacaggttctctactgtttgtgatctctctccccaacagatgggca 1430  
QY 1286 ttgaagcttatgtcacaagtgacagctgttccagagaaaccccaactacgatatgaaata 1345  
|||||  
Db 1431 ttgaagcttatgtcacaagtgacagctgttccagagaaaccccaactacgatatgaaata 1490  
QY 1346 atgaagaagcggagaacatgatgatgatcttactgtatcttgaatgtgagtgtcaggt 1405  
|||||  
Db 1491 atgaagaagcggagaacatgatgatgatcttactgtatcttgaatgtgagtgtcaggt 1550  
QY 1406 ttgatgatgaacatctctccttctcttcttcccaaatctgcgcagttgtccagaagatcccta 1465  
|||||  
Db 1551 ttgatgatgaacatctctccttctcttcttcccaaatctgcgcagttgtccagaagatcccta 1610  
QY 1466 aaacttggttacctatcatctgtcgtggaagagagactgtgccttcccttagtcc 1525  
|||||  
Db 1611 aaacttggttacctatcatctgtcgtggaagagagactgtgccttcccttagtcc 1670  
QY 1526 tgcgcccgatgacagaagttataaaagtcacaatatttgaaacatggcctcagcggatgt 1585  
|||||  
Db 1671 tgcgcccgatgacagaagttataaaagtcacaatatttgaaacatggcctcagcggatgt 1730  
QY 1586 gtagaagaatcaaaaagtcggattatgtgcatcaacagatgaacactttaaagctcgtg 1645  
|||||  
Db 1731 gtagaagaatcaaaaagtcggattatgtgcatcaacagatgaacactttaaagctcgtg 1790  
QY 1646 aagcttatcagcatgaatcaagaatcttgggaccttacttattatggggaagttggagaa 1705  
|||||  
Db 1791 aagcttatcagcatgaatcaagaatcttgggaccttacttattatggggaagttggagaa 1850  
QY 1706 caactgttatattatgaagaatcaagcaagcagacacataacatctaacctcagcga 1765  
|||||  
Db 1851 caactgttatattatgaagaatcaagcaagcagacacataacatctaacctcagcga 1910

|    |      |   |      |
|----|------|---|------|
| QY | 1766 | tccctgatgctccgctccctctgttatctcaaggagataccaaaaggtgtataaaatttgaaag    | 1825 |
| Db | 1911 | tccctgatgctccgctccctctgttatctcaaggagattaccaaaaggtgtataaaacatttgaaag | 1970 |
| QY | 1826 | atttcccaattctgcgcgggggaatatctcaaatataatbtagacgttaccgttagtaagaatg    | 1885 |
| Db | 1971 | atttcccaattctgcgcgggggaatatctcaaatataatbtagacgttaccgttagtaagaatg    | 2030 |
| QY | 1886 | ggccaactaaatcatagatcctcgtgtcccttgaaoccgctatactctagtcttcttaatbtg     | 1945 |
| Db | 2031 | ggccaactaaatcatagatcctcgtgtcccttgaaoccgctatactctagtcttcttaatbtg     | 2090 |
| QY | 1946 | agagaaatcttagttctcaagctcaatttggccctctccccaactgtgtataaagaatctgtag    | 2005 |
| Db | 2091 | agagaaatcttagttctcaagctcaatttggccctctccccaactgtgtataaagaatctgtag    | 2150 |
| QY | 2006 | atcaaaagaagaaacccagataatgtcagacagaagaatgttcaatccgtttttctgtatgt      | 2065 |
| Db | 2151 | atcaaaagaagaaacccagataatgtcagacagaagaatgttcaatccgtttttctgtatgt      | 2210 |
| QY | 2066 | atgagaaccggaagcttggttacctcacagaagaatatacacagcttcttcccaatccagctg     | 2125 |
| Db | 2211 | atgagaaccggaagcttggttacctcacagaagaatatacacagcttcttcccaatccagctg     | 2270 |
| QY | 2126 | gagtgcaagcttggagatccagagattccaaagcttcccaacatcatgtcacagaatcaatgtct   | 2185 |
| Db | 2271 | gagtgcaagcttggagatccagagattccaaagcttcccaacatcatgtcacagaatcaatgtct   | 2330 |
| QY | 2186 | atgttttttatagtttgcaagtcttcaagtctgtttgtgtgtgaatggaggtgtacgtgtacatc   | 2245 |
| Db | 2331 | atgttttttatagtttgcaagtcttcaagtctgtttgtgtgtgaatggaggtgtacgtgtacatc   | 2380 |
| QY | 2246 | taagcaatttgagacacagacacttccctctctgtctctctctcgtgaatataccttcaaac      | 2305 |
| Db | 2391 | taagcaatttgagacacagacacttccctctctgtctctctctcgtgaatataccttcaaac      | 2450 |
| QY | 2306 | acaaaaatgctctctatgaagacacacactcaacctatcccatctcgaaggaaactgtcttca     | 2365 |
| Db | 2451 | acaaaaatgctctctatgaagacacacactcaacctatcccatctcgaaggaaactgtcttca     | 2510 |
| QY | 2366 | tgtcgtatggaaaaacccaggtctctgtgaattctgggtgtgcacaactctgaactttcgggaca   | 2425 |
| Db | 2511 | tgtcgtatggaaaaacccaggtctctgtgaattctgggtgtgtccaaactctgaactttcgggaca  | 2570 |
| QY | 2426 | ggggcactgacccgctctacttgaaggttctctgtgtgtgaacaagaacacgtgtattatcag     | 2485 |
| Db | 2571 | ggggcactgacccgctctacttgaaggttctctgtgtgtgaacaagaacacgtgtattatcag     | 2630 |
| QY | 2486 | aggaacagttatgtaaagatatcttcagcatatctgtctgtgtataaacaatgtccaatgtga     | 2545 |
| Db | 2631 | aggaacagttatgtaaagatatcttcagcatatctgtctgtgtataaacaatgtccaatgtga     | 2690 |
| QY | 2546 | gaagctctccccaagaattctcaagcacacctagcacttagcgaagaacatttatgtccacc      | 2605 |
| Db | 2691 | gaagctctccccaagaattctcaagcacacctagcacttagcgaagaacatttatgtccacc      | 2708 |
| QY | 2606 | caaccagttcttgaaaccccatcaacgggaaataaactcgttaactctctcagttcagatcaag    | 2665 |
| Db | 2709 | caaccagttcttgaaaccccatcaacgggaaataaactcgttaactctctcagttcagatcaag    | 2768 |
| QY | 2666 | aggaaaattgactatgtatgatacatatcagttgaanaatgaagaagaatttgcatt           | 2725 |
| Db | 2769 | aggaaaattgactatgtatgatacatatcagttgaanaatgaagaagaatttgcatt           | 2828 |
| QY | 2726 | atgatgaggaatgaanaatcagagagcccccgcagctttcaaaagaanaacgcgacatattta     | 2785 |
| Db | 2829 | atgatgaggaatgaanaatcagagagcccccgcagctttcaaaagaanaacgcgacatattta     | 2888 |
| QY | 2786 | tctgcctcagttgggaagctcttggtatataatggatgtgaatccccaacgtttcttaagaa      | 2845 |
| Db | 2889 | tctgcctcagttgggaagctcttggtatataatggatgtgaatccccaacgtttcttaagaa      | 2948 |
| QY | 2846 | acagaggtcacaagttggcagatgtccctctcagttctcaagaagaagtgtttctcagaatctt    | 2905 |

|    |      |   |      |
|----|------|---|------|
| Db | 2949 | acaaggccagagtgagcagtgctccctccagctcaagaagttgtttccaggattctactg        | 3008 |
| Qy | 2906 | atgagctcttcaactcaagcccttaaccgttgagaaactaaatgacaattggaccctcgg        | 2965 |
| Db | 3009 | atgagctcttcaactcaagcccttaaccgttgagaaactaaatgacaattggaccctcgg        | 3068 |
| Qy | 2966 | ggccatatataagagcagaagtgtgaagatataatcattgtaactttcagaatacagcct        | 3023 |
| Db | 3069 | ggccatatataagagcagaagtgtgaagatataatcattgtaactttcagaatacagcct        | 3128 |
| Qy | 3026 | ctcgccctattctcttctatcttctgacttcttcttatgaggaagtcaggaagcagag          | 3085 |
| Db | 3129 | ctcgccctattctcttctatcttctgacttcttcttatgaggaagtcaggaagcagag          | 3188 |
| Qy | 3086 | cagaaacctagaaaaaactcttctcaagccttaatgaaaccaaaccttctcttgaaagc         | 3145 |
| Db | 3189 | cagaaacctagaaaaaactcttctcaagccttaatgaaaccaaaccttctcttgaaagc         | 3248 |
| Qy | 3146 | aacatcataatgagccaccataaagaatgagtttgaactgacgcaagccttgagcttatctctg    | 3205 |
| Db | 3249 | aacatcataatgagccaccataaagaatgagtttgaactgacgcaagccttgagcttatctctg    | 3308 |
| Qy | 3206 | atgttgaccctggaaaaaagatgtgacctccaagccttgatttgaccctctcgtcttcaca       | 3265 |
| Db | 3309 | atgttgaccctggaaaaaagatgtgacctccaagccttgatttgaccctctcgtcttcaca       | 3368 |
| Qy | 3266 | cttaaacactgaacccctgctcatgtggagacaagtgtgacagtagacaggaattctgctgttt    | 3325 |
| Db | 3369 | cttaaacacactgaacccctgctcatgtggagacaagatgtgacagtagacaggaattctgctgttt | 3428 |
| Qy | 3326 | tcaaccatcttgatlgagacccaanaagctgtgacttccactgnaaataatgaaagaagaactgca  | 3385 |
| Db | 3429 | tcaaccatcttgatlgagacccaanaagctgtgacttccactgnaaataatgaaagaagaactgca  | 3488 |
| Qy | 3386 | gggctcccttgcaataatccagatlggaagatcccaacttttaagagaattatcgcttcagt      | 3445 |
| Db | 3489 | gggctcccttgcaataatccagatlggaagatcccaacttttaagagaattatcgcttcagt      | 3548 |
| Qy | 3446 | caatcaatggttatataatgatcatcattcccttgcttaagtaatgggtctcagatccaaga      | 3505 |
| Db | 3549 | caatcaatggtctataataatgatcatcattcccttgcttaagtaatgggtctcagatccaaga    | 3608 |
| Qy | 3506 | ttcgatggtatctgcgtccagcaatggcgagcaatgaaaaacatcatctatctattcagtg       | 3565 |
| Db | 3609 | ttcgatggtatctgcgtccagcaatggcgagcaatgaaaaacatcatctatctattcagtg       | 3668 |
| Qy | 3566 | gacatggttcttaactgctacgtaagaaaaaagagagatataaatgtagactgttacaatctctatc | 3625 |
| Db | 3669 | gacatggttcttaactgctacgtaagaaaaaagagagatataaatgtagactgttacaatctctatc | 3728 |
| Qy | 3626 | cagatggttcttgagacaagtgtgaaatgtgtacacatccaagaactgtgaatttggcggtggat   | 3685 |
| Db | 3729 | cagatggttcttgagacaagtgtgaaatgtgtacacatccaagaactgtgaatttggcggtggat   | 3788 |
| Qy | 3686 | gacctatgtgcgagacatctacatgctctgagatgtgagcaacatttctctgtgtacagcaata    | 3745 |
| Db | 3789 | gacctatgtgcgagacatctacatgctctgagatgtgagcaacatttctctgtgtacagcaata    | 3848 |
| Qy | 3746 | agtgatcagaatcccttgaggaaatggtctctgagacacattagagaatttcaagattacagctt   | 3805 |
| Db | 3849 | agtgatcagaatcccttgaggaaatggtctctgagacacattagagaatttcaagattacagctt   | 3908 |
| Qy | 3806 | cagagaacaatatgagacatatggcccccaagaagtgtgcagacttaattatccggatcaaca     | 3865 |
| Db | 3909 | cagagaacaatatgagacatatggcccccaagaagtgtgcagacttaattatccggatcaaca     | 3968 |
| Qy | 3866 | atgccctgagacacaagaagagcccttctcttgatgtcaagtgagatcgtgttgagaccaatgta   | 3925 |
| Db | 3969 | atgccctgagacacaagaagagcccttctcttgatgtgtcaagtgagatcgtgttgagaccaatgta | 4028 |
| Qy | 3926 | ttatccaaggtcalcaagaaccgaagtgtccgttcaagaagttctccagcctctacatctctc     | 3985 |



|   |      |  |      |
|---|------|--|------|
| Db  | 4029 | ttattcaaggcatcaagaaccacgagtgcccgctcagaagttcttccagctctacaattctctc   | 4088 |
| Oy  | 3986 | agttatcatcatcgtatagctctgtatcgtgggaagaagtcggcagactatcatcaggaattcca  | 4045 |
| Db  | 4089 | agttatcatcatcgtatagctctgtatcgtgggaagaagtcggcagactatcatcaggaattcca  | 4148 |
| Oy  | 4046 | ctggaaccttaatggtctctctttggcaatggtgatcatctgvggataaacacaaattctc      | 4105 |
| Db  | 4149 | ctggaaccttaatggtctctctttggcaatggtgatcatctgvggataaacacaaattctc      | 4208 |
| Oy  | 4106 | ttaaccttccaatatatgtgcctcgaatactccgtttgcacccaactatattatggtctga      | 4165 |
| Db  | 4209 | ttaaccttccaatatatgtgcctcgaatactccgtttgcacccaactatattatggtctga      | 4268 |
| Oy  | 4166 | gcaactcttcgcatygaattgatctgtgctctgtaatttaaatgattgacgcatactggaa      | 4225 |
| Db  | 4269 | gcaactcttcgcatygaattgatctgtgctctgtaatttaaatgattgacgcatactggaa      | 4328 |
| Oy  | 4226 | tggagagtaagaacaatacatcagatgacagataactctgtctcatcctactttaccaatagt    | 4285 |
| Db  | 4329 | tggagagtaagaacaatacatcagatgacagataactctgtctcatcctactttaccaatagt    | 4388 |
| Oy  | 4286 | ttggaccactgtctctcttcttaaaagctcgcacttccactccaaggagagatgactcctga     | 4345 |
| Db  | 4389 | ttggaccactgtctctcttcttcttaaaagctcgcacttccactccaaggagagatgactcctga  | 4448 |
| Oy  | 4346 | gacctcagtggaataatccaaaagatgtgctcgaatggaattgacttccagaagaactgaag     | 4405 |
| Db  | 4449 | gacctcagtggaataatccaaaagatgtgctcgaatggaattgacttccagaagaactgaag     | 4508 |
| Oy  | 4406 | tcacagaagatgaactactcaggggaataaaatctctgcttaccagcaatgatgtgaagaagt    | 4465 |
| Db  | 4509 | tcacagaagatgaactactcaggggaataaaatctctgcttaccagcaatgatgtgaagaagt    | 4568 |
| Oy  | 4466 | tcgccatcttcacgaactcaagaatgtgacctcagtggaactctcttttttccaagaatgcaag   | 4525 |
| Db  | 4569 | tcgccatcttcacgaactcaagaatgtgacctcagtggaactctcttttttccaagaatgcaag   | 4628 |
| Oy  | 4526 | taaaagtgttttcagggaataatcaagaactctcttaacacctgtgtgaactcttagaccac     | 4585 |
| Db  | 4629 | taaaagtgttttcagggaataatcaagaactctcttaacacctgtgtgaactcttagaccac     | 4688 |
| Oy  | 4586 | cgttactgactcgtctacacttcgaattcaaccctcagagatgtgggtgcacacagattgccttga | 4645 |
| Db  | 4689 | cgttactgactcgtctacacttcgaattcaaccctcagagatgtgggtgcacacagattgccttga | 4748 |
| Oy  | 4646 | ggatggagaggtctcgggctcggagggcagcagacctctactgactcgagc                | 4693 |
| Db  | 4749 | ggatggagaggtctcgggctcggagggcagcagacctctactgactcgagc                | 4796 |
| RESULT 6  |      |  |      |
| US-09-470-618-13  |      |  |      |
| : Sequence 13, Application US/09470618                                      |      |  |      |
| : Patent No. 6200560  |      |  |      |
| : GENERAL INFORMATION:  |      |  |      |
| : APPLICANT: Contosí, Peter B.  |      |  |      |
| : APPLICANT: Contosí, Peter C.  |      |  |      |
| : TITLE OF INVENTION: Adeno-Associated Vectors for Expression of Factor VII |      |  |      |
| : FILE REFERENCE: Avigen-04082  |      |  |      |
| : CURRENT APPLICATION NUMBER: US/09/470, 618                                |      |  |      |
| : EARLIER FILING DATE: 1999-12-22   |      |  |      |
| : EARLIER APPLICATION NUMBER: 09/364, 862                                   |      |  |      |
| : EARLIER FILING DATE: 1999-07-30   |      |  |      |
| : EARLIER APPLICATION NUMBER: 60/125, 974                                   |      |  |      |
| : EARLIER FILING DATE: 1999-03-24   |      |  |      |
| : EARLIER APPLICATION NUMBER: 60/104, 994                                   |      |  |      |
| : EARLIER FILING DATE: 1998-10-20   |      |  |      |
| : NUMBER OF SEQ ID NOS: 15  |      |  |      |
| : SOFTWARE: Patentlin Ver. 2.0  |      |  |      |
| : SEQ ID NO 13  |      |  |      |
| : LENGTH: 11933   |      |  |      |

|                       | TYPE: DNA    | ORGANISM: Artificial Sequence                                       | FEATURE:   | OTHER INFORMATION: Description of Artificial Sequence: Synthetic |
|-----------------------|--------------|---|------------|--|
| Query Match:          | 89.4%;       | Score 4259.4;   | DB 4;      | Length 11933;  |
| Best Local Similarity | 98.2%;       | Pred. No. 0;  |            |  |
| Matches 4348:         | Conservative | 0;  | Mismatches | 1; Indels 78; Gaps 1;  |
| QY                    | 266          | ccacacatgcaaatagagctccacgctctctctctgctcttgcccttgcatctgcttta         | 325        |  |
| Db                    | 422          | ccacacatgcaaatagagctccacgctctctctctgctcttgcccttgcatctgcttta         | 481        |  |
| QY                    | 326          | gtgcccacgaagatgacacacgctggctggcagtgaggaactcaatgggacatctgcaaatg      | 385        |  |
| Db                    | 482          | gtgcccacgaagatgacacacgctggctggcagtgaggaactcaatgggacatctgcaaatg      | 541        |  |
| QY                    | 386          | atctcggtgagctgctctgtagcgaagaattcccccagagtgcccaaatctttccat           | 445        |  |
| Db                    | 542          | atctcggtgagctgctctgtagcgaagaattcccccagagtgcccaaatctttccat           | 601        |  |
| QY                    | 446          | tcaaacactcaagtcgtgtacaataaagactctgtttgtagaattcaacggttcaaccttca      | 505        |  |
| Db                    | 602          | tcaaacactcaagtcgtgtacaataaagactctgtttgtagaattcaacggttcaaccttca      | 661        |  |
| QY                    | 506          | acatctgctcaagcccaagccacccctgtagtggtctgctgagctccatccagcttagg         | 565        |  |
| Db                    | 662          | acatctgctcaagcccaagccacccctgtagtggtctgctgagctccatccagcttagg         | 721        |  |
| QY                    | 566          | ttatagatcacagtggtgcatcaacttaagaacatggtcccatctgtaactcttaag           | 625        |  |
| Db                    | 722          | ttatagatcacagtggtgcatcaacttaagaacatggtcccatctgtaactcttaag           | 781        |  |
| QY                    | 626          | ctgtgtgtgtatcctactctgcaaaagctctcttgagggagctgtaatatgatatcagaacagtc   | 685        |  |
| Db                    | 782          | ctgtgtgtgtatcctactctgcaaaagctctcttgagggagctgtaatatgatatcagaacagtc   | 841        |  |
| QY                    | 686          | aaagggaaagaagaatgataataagctctccctgctgtaggaagccatatactgtctgcagg      | 745        |  |
| Db                    | 842          | aaagggaaagaagaatgataataagctctccctgctgtaggaagccatatactgtctgcagg      | 901        |  |
| QY                    | 746          | tccttaagaagaatgtgtccaatgctctgaccccttgagccacttgcttaaccatactctt       | 805        |  |
| Db                    | 902          | tccttaagaagaatgtgtccaatgctctgaccccttgagccacttgcttaaccatactctt       | 961        |  |
| QY                    | 806          | ctcatgtgagactctgttaaaagaacttgtaattcaagggcctaatcttgagagcccttaagtatga | 865        |  |
| Db                    | 962          | ctcatgtgagactctgttaaaagaacttgtaattcaagggcctaatcttgagagcccttaagtatga | 1021       |  |
| QY                    | 866          | gagaaaggagctctggtccaaggaagaagacacagacctgtgcacaaattatactactctttg     | 925        |  |
| Db                    | 1022         | gagaaaggagctctggtccaaggaagaagacacagacctgtgcacaaattatactactctttg     | 1081       |  |
| QY                    | 926          | ctgtatctgtatgaaagggaagaagctgtgacactgaagaacaaagaacctccttgatggaagga   | 985        |  |
| Db                    | 1082         | ctgtatctgtatgaaagggaagaagctgtgacactgaagaacaaagaacctccttgatggaagga   | 1141       |  |
| QY                    | 986          | gggagtgcgcatctctctgcgggctctgaccttaaatgcaacagctcaatlggtatgtaaa       | 1045       |  |
| Db                    | 1142         | gggagtgcgcatctctctgcgggctctgaccttaaatgcaacagctcaatlggtatgtaaa       | 1201       |  |
| QY                    | 1046         | ggctctctgcaggtctgattgtagatgccaaggaatacagctctatcttgagatgtagaa        | 1105       |  |
| Db                    | 1202         | ggctctctgcaggtctgattgtagatgccaaggaatacagctctatcttgagatgtagaa        | 1261       |  |
| QY                    | 1106         | tgggagacaacccctggaatgtagctccaataattccctcggaagctcaacaattctcttgagg    | 1165       |  |
| Db                    | 1262         | tgggagacaacccctggaatgtagctccaataattccctcggaagctcaacaattctcttgagg    | 1321       |  |
| QY                    | 1166         | acacatgcgcaaggctctcttggaatactctgcacataaacttctctactgctcaaacacct      | 1225       |  |

Db 1322 accatcgccaggcgtcccttggnaatctcgccaatacttcccttaactgctcaaacactct 1381  
QY 1226 tgatggaccttggagcagttcttactggttttgcatactcttccccaacaaatgata 1285  
Db 1382 tgatggaccttggagcagttcttactggttttgcatactcttccccaacaaatgata 1441  
QY 1286 tggaaagcttalytcaaaatgagacagctgtccagaggnaacccccaactagaaatgaaaaa 1345  
Db 1442 tggaaagcttalytcaaaatgagacagctgtccagaggnaacccccaactagaaatgaaaaa 1501  
QY 1346 atgaagaagcggaaagactatgatgtatcttactgattcttggaaatggaatgtgtcgaagt 1405  
Db 1502 atgaagaagcggaaagactatgatgtatcttactgattcttggaaatggaatgtgtcgaagt 1561  
QY 1406 ttgatagtgaacactctccctcccttaaccgaatcgctcagtttccagaagagctccca 1465  
Db 1562 ttgatagtgaacactctccctcccttaaccgaatcgctcagtttccagaagagctccca 1621  
QY 1466 aaacttgggtacatlaacatgtcgtctgaagaaggagacttgggactatgctccctatgctc 1525  
Db 1622 aaacttgggtacatlaacatgtcgtctgaagaaggagacttgggactatgctccctatgctc 1681  
QY 1526 tcgccccgatgacagaagatataaaatgcaatatttgaacaatgtgcctcagcgagattg 1585  
Db 1682 tcgccccgatgacagaagatataaaatgcaatatttgaacaatgtgcctcagcgagattg 1741  
QY 1586 gtaggaagtacaacaaagctccgattatgtgcatacacagatgaaaccccttaagactctg 1645  
Db 1742 gtaggaagtacaacaaagctccgattatgtgcatacacagatgaaaccccttaagactctg 1801  
QY 1646 aagctatcagcatgaaatcaggaatcttggagaccttacttataatggggaaatgtgagaca 1705  
Db 1802 aagctatcagcatgaaatcaggaatcttggagaccttacttataatggggaaatgtgagaca 1861  
QY 1706 cactgttgcattatatttcaagaatacagaacagacacataaacctcaccctccagaa 1765  
Db 1862 cactgttgcattatatttcaagaatacagaacagacacataaacctcaccctccagaa 1921  
QY 1766 tcaactgtctcgcctctgttattccaaggagattaccaaaaggtgtaaacatttggagg 1825  
Db 1922 tcaactgtctcgcctctgttattccaaggagattaccaaaaggtgtgtaaacatttggagg 1981  
QY 1826 atttccaattctgcagaggagaaatatcaaatataaaatggacagtgacgttgaagaatg 1885  
Db 1982 atttccaattctgcagaggagaaatatcaaatataaaatggacagtgacgttgaagaatg 2041  
QY 1886 ggccaataataatcagatccctgtgtgcctgagccgactatctactctagtttcgttaata 1945  
Db 2042 ggccaataataatcagatccctgtgtgcctgagccgactatctactctagtttcgttaata 2101  
QY 1946 agagaagatcagcttccagactcatgtgcctccctcctcatctgtctacaagaatctgtag 2005  
Db 2102 agagaagatcagcttccagactcatgtgcctccctcctcatctgtctacaagaatctgtag 2161  
QY 2006 atcaaaaggagaaacagataaatgttcaagaagaaggaaatgtcatccgttcttcgtatatt 2065  
Db 2162 atcaaaaggagaaacagataaatgttcaagaagaaggaaatgtcatccgttcttcgtatatt 2221  
QY 2066 atgaagaagcgaagctgtgtactcctcagagaaatatacaacgcttctccccaatccagctg 2125  
Db 2222 atgaagaagcgaagctgtgtactcctcagagaaatatacaacgcttctccccaatccagctg 2281  
QY 2126 gaatgcagcttggagatccagagttccaagctccaacatcatgacagacataatgtcct 2185  
Db 2282 gaatgcagcttggagatccagagttccaagctccaacatcatgacagacataatgtcct 2341  
QY 2186 atgttttttgaatgttgcagttgtcagatttgttgcataatgggtggacatactgtacattc 2245  
Db 2342 atgttttttgaatgttgcagttgtcagatttgttgcataatgggtggacatactgtacattc 2401  
QY 2246 taagcatctgagacagacagactcctcttctgtcctctctcctcgtgatatcactccaac 2305  
Db 2402 taagcatctgagacagacagactcctcttctgtcctctctcctcgtgatatcactccaac 2461

QY 2306 acaaaatgtctatgaagacacacactcacctattcccatcttcagaggaaactgtctca 2365  
Db 2462 acaaaatgtctatgaagacacacactcacctattcccatcttcagaggaaactgtctca 2521  
QY 2366 tgcgatacgaaaacccaggtctatggaattctgtgggtgcaccaactcagacttctggaca 2425  
Db 2522 tgcgatacgaaaacccaggtctatggaattctgtgggtgcaccaactcagacttctggaca 2581  
QY 2426 gaggatagccgccttactgaaggttttactgttgcacaaacacactggtgattatcag 2485  
Db 2582 gaggatagccgccttactgaaggttttactgttgcacaaacacactggtgattatcag 2641  
QY 2486 aggaagattatgaagataattcagatatactgtcgtgagttaaaacaatgtacattgaacaa 2545  
Db 2642 aggaagattatgaagataattcagatatactgtcgtgagttaaaacaatgtacattgaacaa 2701  
QY 2546 gaagctctcccaagaattccaagacacctagcactaggcaaaagcaatttaatgcaccc 2605  
Db 2702 gaagctctcccaagaattccaagacacctagcactaggcaaaagcaatttaatgcaccc 2709  
QY 2606 caccagctcttgaagaagcacaagcgaataaactcgtactactcttcagtacagataag 2665  
Db 2710 caccagctcttgaagaagcacaagcgaataaactcgtactactcttcagtacagataag 2743  
QY 2666 aggaattgactatgatatacatatcagttgaaatgaagaagaaatlttgacattt 2725  
Db 2744 aggaattgactatgatatacatatcagttgaaatgaagaagaaatlttgacattt 2803  
QY 2746 atgatgaagatgaaaaatcaagagcccgagcttcaaaaagaacacagacataattta 2785  
Db 2804 atgatgaagatgaaaaatcaagagcccgagcttcaaaaagaacacagacataattta 2863  
QY 2786 ttgcgcagtgagagaggtcgtgggattatggagatgagttccccaatggttttaagaa 2845  
Db 2864 ttgcgcagtgagagaggtcgtgggattatggagatgagttccccaatggttttaagaa 2923  
QY 2846 acaagggtccagagttgcaggttccctcaggtccaagaaagttgtttccaggaatttactg 2905  
Db 2924 acaagggtccagagttgcaggttccctcaggtccaagaaagttgtttccaggaatttactg 2983  
QY 2906 atggtccttactacagcccttataacggttgagaactaaatlgaaacatttgggactcgtg 2965  
Db 2984 atggtccttactacagcccttataacggttgagaactaaatlgaaacatttgggactcgtg 3043  
QY 2966 ggccatataagagcgaagaattgaagataatcatatgtaacttgaagaatcagaagcct 3025  
Db 3044 ggccatataagagcgaagaattgaagataatcatatgtaacttgaagaatcagaagcct 3103  
QY 3026 ctgcgtccctatctccttactatctagccttatttcttaatggagaagatcagaaggcagag 3085  
Db 3104 ctgcgtccctatctccttactatctagccttatttcttaatggagaagatcagaaggcagag 3163  
QY 3086 cagaacctagaacaaacttgcagaagcctaagaaacccaacttactcttggaaagtgc 3145  
Db 3164 cagaacctagaacaaacttgcagaagcctaagaaacccaacttactcttggaaagtgc 3223  
QY 3146 aacatcataatggcaccacctaagaatgatttgaacttccaagccctgggcttatttctctg 3205  
Db 3224 aacatcataatggcaccacctaagaatgatttgaacttccaagccctgggcttatttctctg 3283  
QY 3206 atgttgaccttgaaaagaatgtgcaactcagagcctgaaatggacccctctgtctgcacaa 3265  
Db 3284 atgttgaccttgaaaagaatgtgcaactcagagcctgaaatggacccctctgtctgcacaa 3343  
QY 3266 ctaaacacactgaacctgtctcaatggagacaaagtgaacagtacagaaattgtctgtttt 3325  
Db 3344 ctaaacacactgaacctgtctcaatggagacaaagtgaacagtacagaaattgtctgtttt 3403  
QY 3326 tcacatcttcttgatggagacccaaaagctgtgtacttccacttgaanaatatggaaagaactgca 3385  
Db 3404 tcacatcttcttgatggagacccaaaagctgtgtacttccacttgaanaatatggaaagaactgca 3463

QY 3386 gggtctccgtgaatacagaatggagaatccactttaagaagaatattcgttccatg 3445  
|||||  
Db 3464 gggtcccttgataatacagaatggagaatccactttaagaagaatattcgttccatg 3523  
QY 3446 caatcaatggtataataatgatacactcgttctagtaaatggtccagatcaaga 3505  
|||||  
Db 3524 caatcaatggtataataatgatacactcgttctagtaaatggtccagatcaaga 3583  
QY 3506 ttgatgtatctgtctagacgttggtcagcaatggaaaacatccatcttcatcttg 3565  
|||||  
Db 3584 ttgatgtatctgtctagacgttggtcagcaatggaaaacatccatcttcatcttg 3643  
QY 3566 gacatgttctacgttacaagaataaagaagagtaataatgacgttacaatctc 3625  
|||||  
Db 3644 gacatgttctacgttacaagaataaagaagagtaataatgacgttacaatctc 3703  
QY 3626 caagtgcttttgagaagtggaatgttacatccaaaagctggaatttggcgg 3685  
|||||  
Db 3704 caagtgcttttgagaagtggaatgttacatccaaaagctggaatttggcgg 3763  
QY 3686 gcttatttggcagacatctacatgtctggaatgagcacatcttctgttgc 3745  
|||||  
Db 3764 gcttatttggcagacatctacatgtctggaatgagcacatcttctgttgc 3823  
QY 3746 agtctcagactcccttggaatggtctctgacacatagagatttcagatcac 3805  
|||||  
Db 3824 agtctcagactcccttggaatggtctctgacacatagagatttcagatcac 3883  
QY 3806 caagacaatattggaagtggtggtcccaagcttgcagacttcatctccg 3865  
|||||  
Db 3884 caagacaatattggaagtggtggtcccaagcttgcagacttcatctccg 3943  
QY 3866 atgctctggaacacaaagagcccttctctggaatgagtggtcttggcaca 3925  
|||||  
Db 3944 atgctctggaacacaaagagcccttctctggaatgagtggtcttggcaca 4003  
QY 3926 ttattcagacatcaagaagcagtggtccctcagaagctctccagcttaca 3985  
|||||  
Db 4004 ttattcagacatcaagaagcagtggtccctcagaagctctccagcttaca 4063  
QY 3986 agttatcatcatgtatagctctggaatggaagtggtcagacttaccag 4045  
|||||  
Db 4064 agttatcatcatgtatagctctggaatggaagtggtcagacttaccag 4123  
QY 4046 ctggaaccttaatggtcttcttggcaatgtgatcatcttggaataaacaat 4105  
|||||  
Db 4124 ctggaaccttaatggtcttcttggcaatgtgatcatcttggaataaacaat 4183  
QY 4106 ttacacctcaattatgtctcgatatacctgttgcaaccaactcatatagca 4165  
|||||  
Db 4184 ttacacctcaattatgtctcgatatacctgttgcaaccaactcatatagca 4243  
QY 4166 gcaactctcgatgagtgatggtggtctgttaataatgttgcaagtgcaat 4225  
|||||  
Db 4244 gcaactctcgatgagtgatggtggtctgttaataatgttgcaagtgcaat 4303  
QY 4226 ttgagagtaagaataatgagtggaagattactgttcatctacttaacaata 4285  
|||||  
Db 4304 ttgagagtaagaataatgagtggaagattactgttcatctacttaacaata 4363  
QY 4286 ttgcacctgtctccttcaaaagctcgacttcaacctcaaggagagtaatg 4345  
|||||  
Db 4364 ttgcacctgtctccttcaaaagctcgacttcaacctcaaggagagtaatg 4423  
QY 4346 gacctcagtgatataatccaaaagtggtgcaagtggaattccagaagaat 4405  
|||||  
Db 4424 gacctcagtgatataatccaaaagtggtgcaagtggaattccagaagaat 4483  
QY 4406 ttacagagtaaacactcagagtggaatctctgtcttaccagcatgtatgta 4465  
|||||  
Db 4484 ttacagagtaaacactcagagtggaatctctgtcttaccagcatgtatgta 4543  
QY 4466 ttctcatctccagcatgcaagtggtccatcagtgactcttcttccagatg 4525

Db 4544 ttctcatctccagcatgcaagtggtccatcagtgactcttcttccagatgca 4603  
QY 4526 taaagtttttcaagggaatacaaacactcttcaaaccttggtggaatctct 4585  
|||||  
Db 4604 taaagtttttcaagggaatacaaacactcttcaaaccttggtggaatctct 4663  
QY 4586 cgttaccgactcgtctacttccgaattcaccccaagtttggtgacacagatg 4645  
|||||  
Db 4664 cgttaccgactcgtctacttccgaattcaccccaagtttggtgacacagatg 4723  
QY 4646 ggaatgaggtcttggtctggaagcagagccttactgactcgag 4692  
|||||  
Db 4724 ggaatgaggtcttggtctggaagcagagccttactgactcgag 4770

RESULT 7  
US-09-364-862-13  
; Sequence 13, Application US/09364862  
; Patent No. 6221349  
; GENERAL INFORMATION:  
; APPLICANT: Coulo, Linda B.  
; APPLICANT: Colosi, Peter C.  
; TITLE OF INVENTION: ADENO-ASSOCIATED VECTORS FOR EXPRESSION OF FACTOR VIII  
; TITLE OF INVENTION: BY TARGET  
; TITLE OF INVENTION: CELLS  
; FILE REFERENCE: AVIGEN-03743  
; CURRENT APPLICATION NUMBER: US/09/364, 862  
; EARLIER FILING DATE: 1999-07-30  
; EARLIER APPLICATION NUMBER: 60/125,974  
; EARLIER FILING DATE: 1999-03-24  
; EARLIER APPLICATION NUMBER: 60/104,994  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 11933  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-364-862-13

Query Match 89.4%; Score 4259.4; DB 4; Length 11933;  
Best Local Similarity 98.2%; Pred. No. 0;  
Matches 4348; Conservative 0; Mismatches 1; Indels 78; Gaps 1;  
QY 266 ccacatgacaataagagctctccacgtctcttcttctgttgccttggatctgctta 325  
|||||  
Db 422 ccaacatgacaataagagctctccacgtctcttcttctgttgccttggatctgctta 481  
QY 326 gtgcacacagaataactactcgtgtgagtggaatcgtcatcagatataatgca 385  
|||||  
Db 482 gtgcacacagaataactactcgtgtgagtggaatcgtcatcagatataatgca 541  
QY 386 atctcgtgtgagctgctgtgtgagcgaagatttccctcctaagtgccaatcttccat 445  
|||||  
Db 542 atctcgtgtgagctgctgtgtgagcgaagatttccctcctaagtgccaatcttccat 601  
QY 446 tcaaacactcagtcgttataaaaagactgtgttgtgagaattcaaggttcacactttca 505  
|||||  
Db 602 tcaaacactcagtcgttataaaaagactgtgttgtgagaattcaaggttcacactttca 661  
QY 506 acatcgttaagcagaagccacccctgtgagtggtctgtataggcttccatccagctgag 565  
|||||  
Db 662 acatcgttaagcagaagccacccctgtgagtggtctgtataggcttccatccagctgag 721  
QY 566 ttatgatataagtggtatataacttaagaagaatggtctccatctgtatgctcatg 625  
|||||  
Db 722 ttatgatataagtggtatataacttaagaagaatggtctccatctgtatgctcatg 781  
QY 626 ctgtgtgtgtatctacttgtaaaagctcttgagtgagtgatataatgatacagacagctc 685

|||||  
Db 782 ctgttgtgtatcctactggaagcttctgagggagcgtgaatgatgatcagacagctc 841  
QY 686 aaaaagggaagaagaatgaataaaagcttcctccgtgfggaagccaatacatgtctggcagg 745  
Db 842 aaaaagggaagaagaatgaataaaagcttcctccgtgfggaagccaatacatgtctggcagg 901  
QY 746 tccggaagaagaatggtcccaatggcctctgacccactgtgaccttaccatactctt 805  
Db 902 tccggaagaagaatggtcccaatggcctctgacccactgtgaccttaccatactctt 961  
QY 806 ctcatgtggaacctgtgtaaaagaacttgaatccaagcctcatgtgagccctactagtatgta 865  
Db 962 ctcatgtggaacctgtgtaaaagaacttgaatccaagcctcatgtgagccctactagtatgta 1021  
QY 866 gagaagaagaatctgccaagaagaagaagacagaccttgacaatatcttactacttttg 925  
Db 1022 gagaagaagaatctgccaagaagaagaagacagaccttgacaatatcttactacttttg 1081  
QY 926 ctgtatttgtaagaagggaanaagtttggcactcagaacaagaactccttgatgcagata 985  
Db 1082 ctgtatttgtaagaagggaanaagtttggcactcagaacaagaactccttgatgcagata 1141  
QY 986 ggagtgctgcatctgtctgggccttggcctaataatgacacagtcgaatgttataaaca 1045  
Db 1142 ggagtgctgcatctgtctgggccttggcctaataatgacacagtcgaatgttataaaca 1201  
QY 1046 ggtcttcgccaaggtctgaattgataatgccaagaatacagtcatttgatgagatgttgtaa 1105  
Db 1202 ggtcttcgccaaggtctgaattgataatgccaagaatacagtcatttgatgagatgttgtaa 1261  
QY 1106 tgggacacacactccggaagtgacactcaataatcctcgaagaagtcacacacacttctgtgaaga 1165  
Db 1262 tgggacacacactccggaagtgacactcaataatcctcgaagaagtcacacacacttctgtgaaga 1321  
QY 1166 accttcgccaaggtcctcttgtaaatctcgccaataacttccctactctgctcaaacactc 1225  
Db 1322 accttcgccaaggtcctcttgtaaatctcgccaataacttccctactctgctcaaacactc 1381  
QY 1226 tgaatgacctggaagcttctactgttttgcatactcttccccaacaatgaatggca 1285  
Db 1382 tgaatgacctggaagcttctactgttttgcatactcttccccaacaatgaatggca 1441  
QY 1286 tgggaagcttatgtcaaaagtagacagctgtgccagaggaaccccaactacgaatgtaaaaaa 1345  
Db 1442 tgggaagcttatgtcaaaagtagacagctgtgccagaggaaccccaactacgaatgtaaaaaa 1501  
QY 1346 atggaagaagcgggaagactatgatatgtactctgattctggaatgtgatatgtgtcaggt 1405  
Db 1502 atggaagaagcgggaagactatgatatgtactctgattctggaatgtgatatgtgtcaggt 1561  
QY 1406 ttgtatgtgaacaacttctcctccttatccaatcgctcagttgccaagaagatcccta 1465  
Db 1562 ttgtatgtgaacaacttctcctccttatccaatcgctcagttgccaagaagatcccta 1621  
QY 1466 aaacttgggtacatatactgtctgtaagaagggaagacttgggaactatgtcccttaagtcc 1525  
Db 1622 aaacttgggtacatatactgtctgtaagaagggaagacttgggaactatgtcccttaagtcc 1681  
QY 1526 tcgcccccgatgacagaagtataaaagtcataattgaaacaatgcccctcagcgagattg 1585  
Db 1682 tcgcccccgatgacagaagtataaaagtcataattgaaacaatgcccctcagcgagattg 1741  
QY 1586 gtggaagaatcaaaaaagctcgattatgtgcatacacagatgaaaaccttgaagactgtg 1645  
Db 1742 gtggaagaatcaaaaaagctcgattatgtgcatacacagatgaaaaccttgaagactgtg 1801  
QY 1646 aagctatcagcatgtaatcaggaatctctgggaaccttacttataatgggaagtttggaaca 1705  
Db 1802 aagctatcagcatgtaatcaggaatctctgggaaccttacttataatgggaagtttggaaca 1861  
QY 1706 cactgttattatatttaagaatacgaagcagacataaaccttacccttaacgaa 1765  
|||||

Db 1862 cactgttattatatttaagaatacgaagcagacacataaacatactbaccttcacgaa 1921  
QY 1766 tcaactgattccgtctcttctgtatataaaggatataccaaaaggtgtataaacatttgagg 1825  
Db 1922 tcaactgattccgtctcttctgtatataaaggatataccaaaaggtgtataaacatttgagg 1981  
QY 1826 atttccaattctgcagagaagaatatataataatgaatgacagctgttgaagaatg 1885  
Db 1982 atttccaattctgcagagaagaatatataataatgaatgacagctgttgaagaatg 2041  
QY 1886 ggccaactaaatacgaatctctgtgtgcctgaccgctattactagtatttctgtaatatg 1945  
Db 2042 ggccaactaaatacgaatctctgtgtgcctgaccgctattactagtatttctgtaatatg 2101  
QY 1946 agagatctgacttcagagcactatgtgacctccctccctcatctgtgtacaanaaatctgag 2005  
Db 2102 agagatctgacttcagagcactatgtgacctccctccctcatctgtgtacaanaaatctgag 2161  
QY 2006 atcaagaagaacacagataatgttcaagaagaagaatgtcatctgtttctgtatttg 2065  
Db 2162 atcaagaagaacacagataatgttcaagaagaagaatgtcatctgtttctgtatttg 2221  
QY 2066 atgagaacccgaagctgtgtactctacagaagaatatacaagcttcttcccaatccagctg 2125  
Db 2222 atgagaacccgaagctgtgtactctacagaagaatatacaagcttcttcccaatccagctg 2281  
QY 2126 gagtgcagcttgagagatccagatctcagagctcccaacatcatgtcacagcatcaatggtc 2185  
Db 2282 gagtgcagcttgagagatccagatctcagagctcccaacatcatgtcacagcatcaatggtc 2341  
QY 2186 atgttttgataagtttgagttgtaagtttgttgcataagtggaatgtaactgtataatc 2245  
Db 2342 atgttttgataagtttgagttgtaagtttgttgcataagtggaatgtaactgtataatc 2401  
QY 2246 taagatctggagacagaagcttactccttctgtctcttctgtctgtgatatataccttcaac 2305  
Db 2402 taagatctggagacagaagcttactccttctgtctcttctgtctgtgatatataccttcaac 2461  
QY 2306 acaaaatgtctatataaagaacaacatacactcactatccacttcacaggaagaactgtctca 2365  
Db 2462 acaaaatgtctatataaagaacaacatacactcactatccacttcacaggaagaactgtctca 2521  
QY 2366 tgtcgaatggaaaacccaggtctatagatctctgggtgtcccaactcagacttctcggaaca 2425  
Db 2522 tgtcgaatggaaaacccaggtctatagatctctgggtgtcccaactcagacttctcggaaca 2581  
QY 2426 gaggcataccgcttactgaaggttttctagtgttgaanaagaacatggtgattatgag 2485  
Db 2582 gaggcataccgcttactgaaggttttctagtgttgaanaagaacatggtgattatgag 2641  
QY 2486 aggaagttatgaaagataatccaacatactgtcgtgagtaaaaaacaatgccaatgtaacaa 2545  
Db 2642 aggaagttatgaaagataatccaacatactgtcgtgagtaaaaaacaatgccaatgtaacaa 2701  
QY 2546 gaagcttctccagaatccaagacacccctagcactaggaagcaaatctaattgccaacc 2605  
Db 2702 gaagcttctccagaatccaagacacccctagcactaggaagcaaatctaattgccaacc 2709  
QY 2606 caccagctcttgaaaacgcatacaggaataaactgtactactctcagtcagatcaag 2665  
Db 2710 -----gaaataactctgtactactccttcagtcagatcaag 2743  
QY 2666 aggaatctgactatgatataccatatcaagtctggaatgaaagaagaatttgaactt 2725  
Db 2744 aggaatctgactatgatataccatatcaagtctggaatgaaagaagaatttgaactt 2803  
QY 2726 atgatgagatgaaataatcagaagcccccagcttccaagaagaacacgacactatttta 2785  
Db 2804 atgatgagatgaaataatcagaagcccccagcttccaagaagaacacgacactatttta 2863  
QY 2786 ttgctgcagtggaagagctcgtggaatataatggatgaatgaatgctcccaatggtttaaaga 2845  
Db 2864 ttgctgcagtggaagagctcgtggaatataatggatgaatgaatgctcccaatggtttaaaga 2923  
|||||



```

1 CURRENT APPLICATION DATA:
2 APPLICATION NUMBER: US/08/882,083
3 FILING DATE:
4 CLASSIFICATION: 514
5 PRIOR APPLICATION DATA:
6 APPLICATION NUMBER: US 08/558,107
7 FILING DATE: 13-NOV-1995
8 ATTORNEY/AGENT INFORMATION:
9 NAME: ISACSON, John P.
10 REGISTRATION NUMBER: 33,715
11 REFERENCE/DOCKET NUMBER: 30472/212
12 TELECOMMUNICATION INFORMATION:
13 TELEPHONE: (202)672-5300
14 TELEFAX: (202)672-5399
15 TELEX: 904136
16 INFORMATION FOR SEQ ID NO: 1:
17 SEQUENCE CHARACTERISTICS:
18 LENGTH: 5035 base pairs
19 TYPE: nucleic acid
20 STRANDEDNESS: single
21 TOPOLOGY: linear
22 FEATURE:
23 NAME/KEY: CDS
24 LOCATION: 35..5017
25 US-08-882-083-1

```

|                            |        |                |             |              |
|----------------------------|--------|----------------|-------------|--------------|
| Query Match                | 78.78; | Score 3748;    | DB 2;       | Length 5035; |
| Best Local Similarity      | 87.48; | Pred. No. 0;   |             |              |
| Matches 4374; Conservative | 0;     | Mismatches 60; | Indels 570; | Gaps 2       |

[illegible]

|    |      |  |      |
|----|------|--|------|
| Db | 630  | GAGAAAGGAGTCGTGGCCACAGSAAAGACACAGACCTTGCACAAATTTATACACTTTTGG     | 689  |
| Qy | 926  | ctgtatcttgatgtaeagggaaaagttgacactcagaaacaagaacctcccttgatgacgata  | 985  |
| Db | 690  | CTGTATTTTGATGAAGGGGAAAAAGTTGGCACACCAAAAGAACTCCTTGATGACGAGATA     | 749  |
| Qy | 986  | ggagatcgacatcgctcccgaggcccgagcctaaatgcaacagacaaatggtatgtataaca   | 1045 |
| Db | 750  | GGGAGTCTGCACTCTGCTGGGCGCTGAATATGACACACAGTCATATGGTTATATGTAACA     | 809  |
| Qy | 1046 | ggtctctgcccagtcctgatlgatgtaagcacaaggaaatcagctctatctgcaatgtaatlga | 1105 |
| Db | 810  | GGTCTCTCCAGGTCGTGATTTGATGGCCACAGAAATACGTATTGTGCATGTGATTTGGA      | 869  |
| Qy | 1106 | tgggacacacatccctgaagtgacatcaatctctcogaagtgacacatctcttgtagga      | 1165 |
| Db | 870  | TGGGCAACCACTCTGAAGTGCACCATATTTCCGAAAGTGACACATTTCTTGTAGAGA        | 929  |
| Qy | 1166 | acacatcgccagtgctcttggaatctcgccaatacttccctactgctgaacacct          | 1225 |
| Db | 930  | ACCATCGCCAGGCGTCCTGTGAAATTCGCGCAATACCTTTCTTACTGTGTCAACACCT       | 989  |
| Qy | 1226 | tgatgacctggagacgtttctactgtttctgtcatactctccccaacaatgaatgga        | 1285 |
| Db | 990  | TGATGACCTTGGACAGTTTCTACTGTTTTGTATATCTTTCGCCAACATGTATGGCA         | 1049 |
| Qy | 1286 | tggaagcttatgtlcaaaagtagacagctgtlccagaaggaaacccaactacgaatgtaaaaa  | 1345 |
| Db | 1050 | TGCAAGCTTATGTCAAGTATGACAGCTGTCCAGAGAACCCCACTACGAATGAATAA         | 1109 |
| Qy | 1346 | atgaagaagcggaagactatgatatgatacttactgattctgaaatgtagtgtgctcag      | 1405 |
| Db | 1110 | ATGAAGAAGCGGAAGACTATGATGATGATCTTCTATTTCTGAATGATGTGTGTCAG         | 1169 |
| Qy | 1406 | ttgatgatgaanaactctccctcccttataccaatctgcgcagttgccaagaagactcta     | 1465 |
| Db | 1170 | TTGATGATGACACACTCTCTCTTCTCTTATTCCAATTTGCTCTCACTGGCCAAAGCATCT     | 1229 |
| Qy | 1466 | aaacttggtaacatlaactatgctgctgctgaaggaaggagctggagactatgctccctagtc  | 1525 |
| Db | 1230 | AAACTTGGGTACATTATACATTGTGCTGAAGAGAGAGACTGGGACATGTGCTCCTTA        | 1289 |
| Qy | 1526 | tcgcccccgatgaaagaagtataaaagtcaaatatttgaaacaatgagccctcaagagatt    | 1585 |
| Db | 1290 | TGCGCCCCGATGACAGAAAGTTATTAAGTAAATTTTGAACAATGGCCCTCAGCGATTG       | 1349 |
| Qy | 1586 | gtgaagagtaacaaaagctccgatcttaatgagcaaacagatgaaccttgaagctg         | 1645 |
| Db | 1350 | GTAGGAAGTACAAAAAAGTCCGATTTATGGCATACACAGTGAACCTTTAAGACTCGT        | 1409 |
| Qy | 1646 | aagcatatcagcatgaaatcagaagactcttggaaccttacttlatggygaaglttgagaca   | 1705 |
| Db | 1410 | AAGCATTCACCATGAMTCAGAAATCTTGGACCTTACTTTATGGGAAGTTGAGACA          | 1469 |
| Qy | 1706 | cactgttgattatatttaagaatcaagaagaagacaataatacatctacctcaaggaa       | 1765 |
| Db | 1470 | CACGTGTGATTTATTTTAAAGATCAGCAAGCAGACCATATTAACATCTACCTTCAGGA       | 1529 |
| Qy | 1766 | tcactgtagtccgaccttcttlatccaaggagatcttacaagaagtgtaaaaaatctgaag    | 1825 |
| Db | 1530 | TCACGTAGTCCGTCTTGTATTCAAGGAGATTACCAAAAAGGTGTAATAACATTTGAAG       | 1589 |
| Qy | 1826 | attcttcaactctccagggaaatctcaatataaatggaacagagactgtagaagat         | 1885 |
| Db | 1590 | ATTTTCCATTTCTGCCAGGGAATATTTCAAAATATAATGAGACGTGTAGAAGATG          | 1649 |
| Qy | 1886 | ggccaactaaatcagatccctcggtgacctgaagcccgcatctactctagttcgtltaatag   | 1945 |
| Db | 1650 | GGCCAACTAATCAGATCTCGGGGCTGACCCGCTATTACTCTAGTTTCGTTAATATGG        | 1709 |
| Qy | 1946 | agagagatctagcttcaagacatctggccctctctcatctctgctacaaagaatctgag      | 2005 |
| Db | 1710 | AGAAGAGATACGTTCACGGACTATGGCCCTTCCCTCATCTGTTCACAAAGATCTGGAG       | 1769 |



|    |      |   |      |
|----|------|---|------|
| QY | 2006 | atlaaagagaaacccagataatgtccagaaagagaagtaactccgtgtttctgtatlttg    | 2055 |
| Dp | 1770 | ATCAAAGGAGAAACCAATATATGTCTAGACAAAGAGAAATGTACTCTGTGTTTCTGTATTTG  | 1829 |
| QY | 2066 | atlgaaacccgaagcttgttacctccacagagaataatacaagctttctccccaatccagctg | 2125 |
| Dp | 1830 | ATGGAACCGAAGCTGTACTCTCAGAGAAATATACAAAGCTTTCTCCCAATCCAGCTG       | 1889 |
| QY | 2126 | gaatgcagcttgtgaagatccagagttcccaagcctccaacatcatgcaagcatcaatggct  | 2185 |
| Dp | 1890 | GAGGCAAGCTTGAAGATCCAGAACTTCCAAGCCTCCAACTCATGTCCAGCATCAATGGCT    | 1949 |
| QY | 2186 | atgtttctgaatgttgcattgtcagtttgttgaatgaatggcaataacgctgtacatc      | 2245 |
| Dp | 1950 | ATGTTTGTAGTAGTTTGCAGTTTGTCAAGTTTGTTTGCATGAGGTGGCATACTGTGATTC    | 2009 |
| QY | 2246 | taagcatttgagaacagcagctactccctctctgtctctctctctgtatataccttcaac    | 2305 |
| Dp | 2010 | TAAACATTGGAGACACAGACTGACTTCTTCTGCTCTTCTCTGTGATATACCTTCAAC       | 2069 |
| QY | 2306 | acaaatgtgctcttgaagacaacactaaccttaccattccagtgaaactgtcttca        | 2355 |
| Dp | 2070 | ACAAAATGGTCTATGAAGACACTCACCTCATCTCCATCTTCAGAGAAACTGTCTCA        | 2129 |
| QY | 2366 | ttctgaatgaaaccccaaggtcctatgtatcttgggtgcacaacccaagctttggaca      | 2425 |
| Dp | 2130 | TGTGTATGAAAAACCCAGGCTCTATGTGATCTTG66GTGCCACAACTCAGACTTTGGAAACA  | 2189 |
| QY | 2426 | gaagcatgcagccctacttgaagtttctagtgtgtgacaagaacactggtatattacg      | 2485 |
| Dp | 2190 | GAGGCATAGACCGGCTTACTGAAAGGTTTCAAGTTTATTCAGAGGGGAGAGACGACG       | 2249 |
| QY | 2486 | aggcagcttctgaagatatttca-----gcactgtgtgtaaaaca                   | 2530 |
| Dp | 2250 | ACTTCTGGACCTGGAGAAATATTAGTGAAGACGACGACTACTGCACATCGTGCACA        | 2309 |
| QY | 2531 | atgcatctgaaacaaagaagcttctcccaaatccaagaacccctagcaactagcaaaagc    | 2590 |
| Dp | 2310 | GTCGATTTGAACCAAGAAAGCTTCTCCAGAAATTCAAAGACACCCTAGCCTAGGCAAAAAGC  | 2369 |
| QY | 2591 | aattaatgcac-----  | 2603 |
| Dp | 2370 | AATTATATGCCCCACCAATTCAGAAAAATGACATGAGAGACGTGACCTTGTTGGCAC       | 2429 |
| QY | 2604 | -----   | 2603 |
| Dp | 2490 | GACGAGTCTTACTCCACATGGGCTATCTTATCTGATCTCCAAGAAACCAATATGAGA       | 2549 |
| QY | 2604 | -----   | 2603 |
| Dp | 2550 | CTTTTCTGTATGATCATCATCGTGGAGCAATAGACAGTAATPACAGCCTGTTCAAAATGA    | 2609 |
| QY | 2604 | -----   | 2603 |
| Dp | 2610 | CACACTTCAGGCAAGCTTCATCATCAGTGGGGACATGTAATTAACCTGATCAGGCC        | 2669 |
| QY | 2604 | -----   | 2603 |
| Dp | 2670 | TCCAAATTAGATTAAATGAGAAACTGGGACAATGACAGATCTCTTGCTTGGGATAAC       | 2729 |
| QY | 2604 | -----   | 2603 |
| Dp | 2730 | ACTATGTACTCAGATACCAAAAAGAAAGATGGAATTCACAGAGACGACACCAAAAAAA      | 2789 |
| QY | 2604 | -----   | 2603 |
| Dp | 2790 | CAGCTTTAAGAAAAAGATACATTTTGTCCCTGACGCTTGGAAGCAANTCATGCAA         | 2849 |

[illegible]

Db 3930 AGTATAAAATGGACCTGATCATCTATCCAGGTGTTTTGAGACACATGGAATGTTAC 3989  
3656 catccaaagcttgaaatttgagggttgaaatgcttatgtgagcaatctatagatcttgga 3715  
3990 CATCCAAAGCTGGAATTTGGCGGGTGGAAATGCTTATTTGGGAGCATCTACATGCTGGGA 4049  
3716 tgaacacacttctctgtgtgtacagcaataagtgtagactccctgggaatggctctgt 3775  
4050 TGAGACACACTTTTCTGTGTGTAGACATAATAGTGCAGACTCCCTGGGAATGGCTTGTG 4109  
3776 gacacatagaagatlttcagattacagcttcaagaaacatatgagacgtggcccccaagc 3835  
4110 GACACATATAGAGATTTTCAGATTTACAGCTTCAGACATATGAGACAGTGGGCCCAAGC 4169  
3836 tggccgaacttcatattccgagcaatcaatgtcgtgagacccaagagcccttctt 3895  
4170 TGCCGACACTTCATTATTTCCGATCAATCAATGCTTGAGACCAAGAGCCCTTTTCTT 4229  
3896 ggaatcaagtgatctgtgtgaccaaagtatattcaacgacatcaagacccaaggtgcc 3955  
4230 GGATTCAGAGTGGATCTTGTGGACCAATGATTTCAAGGATCAAGACCCAGGGTGGCC 4289  
3956 gtcaagaagttctccagcctacatctctcagttatcatatgtagtctgtatgga 4015  
4290 GTCAGAAATTCTCCAGCCTCTACATCTCTCAGTTATCATCATGATGATGCTTGTGGA 4349  
4016 agaagtgagcagacttaccgaagaaatccacttggaacttaagtgtctctcttgcaatg 4075  
4350 AGAAGTGGCAGACTTATTCGAGGAATTCACACTGGAACCTTATGCTCTTTTGGCAATG 4409  
4076 tggattcatctggagataaaacacataatttlaacccctcaatattgtcgtatatacc 4135  
4410 TGATTCATCTGGGATTAACACAAATATTTTAAACCTCCCAATATTCCTGATACATCC 4469  
4136 gtttgacccaactcatatagacttgcagcactcttcgtagagttgtgtggctgtg 4195  
4470 GTTTGACCCCAACTCATTTAAGCATTCGACACACTCTTCGATGAGTGTGAGTGGCTGTG 4529  
4196 atttaaatgtgcagatgcacatggaatggagatgaagcaatatcaaatgacaga 4255  
4530 ATTTAATATGTTGACGATGCGCATTTGGGAATGGAGATGAAGCATATACATGACACAGA 4589  
4256 ttaactgctacatccacttaccataatglttgcacactgtgtctctcaaaagctgac 4315  
4590 TTACTGCTCATCTCTACTTACCAATATGTTTGCACCTGGTCTCTTCAAAAGCTGCAC 4649  
4316 ttaactcccaaggagagatgaatgtcctggagactcaaggtgaataatccaaaagtgagc 4375  
4650 TTACCTCCCAAGGAGAGATGATGCTGAGACCTCAAGGTGAATATCAAAAGAGTGGC 4709  
4376 tgcgaatgtgacttccaaagaacaaatgaatgcacagagatgaactactcaaggagtaaat 4435  
4710 TGCAGATGGACTTCAGAAACATGAAGTCAAGAGATGACTACTCAGGAGATTAAT 4769  
4436 ctctgtcttaccagatgtatgtgaaggagttcctcatctccagcagttcaagaatggccalc 4495  
4770 CTCTGCTTACACAGATGATGTGAAGAGTTCCTCATCTCCAGCAGTCAAGATGGCCATC 4829  
4496 agtggagctctcttttccaaatgtgcaaatgaaggttttccagggaatccaagctcct 4555  
4830 AGTGGACTCTCTTTTTCAAATGCAAGTGAAGGTTTTTCAAGGAATCAAGACTCTCT 4889  
4556 tcaacacctgtgtgaactccttagaccacacgcttactgaactcgttacccttgaattcac 4615  
4890 TCACACTGTGTGAATCTCTAGAACCCACCGTTACTGACTCGCTACTCTTGAATTCACC 4949  
4616 cccaagattggtgtgacacagatgtcccttgagagatggaagttctgtggtctgaggaacagg 4675  
4950 CCCAGAGTTGGTGTGACCAAGATTTGCCCTGAGAGATGAGGTTCTGGCTGCGAGGACACAGG 5009  
4676 acctctactgtacctgagaggttc 4699  
|||||

Db 5010 ACTCTACTGAGGGTGGCCACTGC 5033  
RESULT 9  
US-08-558-107-1  
; Sequence 1, Application US/08558107  
; Patent No. 5910481  
; GENERAL INFORMATION:  
; APPLICANT: VOORBERG, Johannes J.  
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/558,107  
; FILING DATE: 13-Nov-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: ISACSON, John P.  
; REGISTRATION NUMBER: 33,715  
; REFERENCE/DOCKET NUMBER: 30472/212  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ. ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5035 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: Linear  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 35..5017  
; US-08-558-107-1  
Query Match 78.7%; Score 3748; DB 2; Length 5035;  
Best Local Similarity 87.4%; Pred. No. 0;  
Matches 4374; Conservative 0; Mismatches 60; Indels 570; Gaps 2;  
266 ccacatgcaaatagaagacttccacactgtcttcttctgtgtgacattcaagttcacccttca 325  
30 CCACCATGGAATATGAGACTCTCCACCTGCTTCTTTCTGTGCTCTTTTGCATTTCTGCTT 89  
326 gtgcaccagaagaatactactgtgtgtgagtggaatgcaatgcaatgcaatgcaatgcaatg 385  
90 GTGCACCAAGAAGATGACTGATGAGTGGTGCAGTGAAGTGCATGATGAGTGAATGCAAGT 149  
386 atctcgatgagctgtcctgtgtgagcaagatcttctccttaagaatgcaaaatcttccat 445  
150 ATCTCGGTGAGCTGCTCTGTGAGCAGCAAGATTTCTCTCTAAGTGCACAAATCTTTTCAAT 209  
446 tcaacacctgaatgtgtgtacaaaagaactgtttgtagaattcaagttcacccttca 505  
210 TCACACTCTGATGTGTGTACAAAAGACTGTGTTGTGAGATTCACAGGATCACCTTTTCA 269  
506 acatcgatgaagccaagccaacctgtgagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 565  
270 ACATCGTGAAGCCAAAGCCACCCCTGATGGTGTGCTAGTGTCTTACATCATCAGGTGAGG 329  
566 ttatgtatacagtggtcatttaacacttaagaacatggtcttccatctctgtcagttctcatg 625  
|||||



Db 330 TTATGATACAGTGTGATATACCTTAAGAACATGGCTTCCATCTGTGAGTCTTCATG 389  
QY 626 ctgtgtgtgtactactactggaagctctcgaaggagctgatatgatgatcaagccgctc 685  
Db 390 CTGTGGGTATCTCTACTGGAAGCTTCTGAGGAGCTGAATGATGATGACACCGTC 449  
QY 686 aaaggagaaagatgatataagctctccctgtgtggaagccatatactgtcggagc 745  
Db 450 AAAGGAGAAAGAGATGATTAAGTCTTCCCTGGTGAAGCCATACATATGTCTGGGAGG 509  
QY 746 tctcgaagaagaatgtgtcccaatgtgcctctgaaccactgtgcctactactacatctt 805  
Db 510 TCCGAAGAAGAAATGCTCCATATGCGCTGTGACCACTGTGCTTACTACTCATATCTTT 569  
QY 806 ctatgtggagacctgtgtaaagacttgatctcaagccctcatgtggagccctactgtatgta 865  
Db 570 CTCTATGTGGACCTGTGTAAAGACTTGAATTCAGGCTCATTTGGGCCCTCTACTATATGTA 629  
QY 866 gaagaaggagctctggccaaggaanaagacacagacctgtgcaanaattatactacttttg 925  
Db 630 GAGAGAGGAGTCTGGCCACAGGAAGACACAGACCTTGCAAAATTTATACTACTTTTTCG 689  
QY 926 ctgtatttgtatgaagggaagaaagtgtgcactcgaagaacaaagaactccttgtatgcaagata 985  
Db 690 CTGTATTTGATGAAGGGGAAAGTTGGCAGCTCAGAAACAAGAACTCCTGTGATGAGGATA 749  
QY 986 gggatgtgatctgtcgtcgggacctgtgctaataatgcacagatcaatagtgtatgtataca 1045  
Db 750 GGGATGTGATCTGTCTGGGAGCTGGGAGCTGCTTAAATGCACAGATCAATGATGATGTAACA 809  
QY 1046 ggtctctgcagagctgtatgtgatgtgcacaggaatcagctcatctgtgcatgtgtatgaa 1105  
Db 810 GGTCTGTGCCAGGTCTGATTTGGATGCCACAGAAATCAGCTTATGGCATGTGATTTGAA 869  
QY 1106 tgggaccacctcctgaagtgacactcaatctcctcgaaggltcaacacatctctgttgagga 1165  
Db 870 TGGGACACCTCTGGAAGTGCACATATTCCTCGAAGGTGCACACTTTCTGTGAGGA 929  
QY 1166 accatgtccagggggtcctctgtgaaatctcgcaataactcttactcgtccaaacactct 1225  
Db 930 ACCATGTCCAGGGGCTCTTGTGAATCTCGCCAAATTAACCTTTACGTGCTAAACACTCT 989  
QY 1226 tgaatgacctctggacagcttctactcgtctgtcatalctctcccaacaacatgatgtga 1285  
Db 990 TGAATGACCTTGGACAGCTTCTACTGTTTGTCTATCTCTTCCACCAATGATGTGCA 1049  
QY 1286 tgggaagctatgtcaaaagtacagagctgtgcagaggaagcccaactacgaatgaanaata 1345  
Db 1050 TGGAACTTATGTCAAAAGTGAACAGCTGTGCAGAGAAACCCAACTACGAATGAANAATA 1109  
QY 1346 atgaagaaggagaaactatgatgatcttactgtatctgaaatgtgatgtgtcaggt 1405  
Db 1110 ATGAAGAAGCGGAAGACTATGATGATGATCTTACTGATTTGAAATGATGTGGTCAAGT 1169  
QY 1406 ttgatgatgacaactcctctctcttataccaaatcgcactcagctgtgccaagaacatccta 1465  
Db 1170 FTGATGATGACAACTCTCTCTCTTATTCCAANTCCCTCAGTTGCCAANAAGCATCTTA 1229  
QY 1466 aaactgtggatcatctacatctgtcgtcgaagaagagagactgtggaactatgtcccttagtc 1525  
Db 1230 AAACCTTGGGTACATTAATGCTGCTGTAAGAGAGAGCACTGGGACTATGCTCCCTTAATGCC 1289  
QY 1526 tgcgccccggatgacagaagaagtataaagctcaatatgtgaacaaatgtgcccccaaggttg 1585  
Db 1290 TCGCCCCCGATGACAGATTATTAAGTCAATATTTGAAACAAATGGCCCTACGCGGATTCG 1349  
QY 1586 gtaggaagtaacaaaagctcgaatttatgtgcatacacagatgaaccccttaagactcgtg 1645  
Db 1350 GTAGGAGATGCAAAAAAGTCCGATTTTATGGCATACACAGATGAACCTTTAAGACTGTG 1409  
QY 1646 aagcatctcagcatgaacgaagactcttggagcccttacttctatgggaagttggagaca 1705  
Db 1410 AAGCTATTCAGCATGAAATCTGGAGATCTTGGACCTTATTTAGGAGAGTTGGAGACA 1469

QY 1706 cactgtgatataatgaagaatcaagcagaccataaacatctaccctcacaggaa 1765  
Db 1470 CACTGTTGATTTATATTAAGAAATCAAGCAAGACCATATATCACTTACCCTACGGAA 1529  
QY 1766 tcaactgtatgcgttcccttltatcaagagatataccaaaggltgtanaacatttgaag 1825  
Db 1530 TCACTGATGTCCGTCCTTTGTATTCAMGAGATTTACCAAAAGGTATTAACATTTGAAG 1589  
QY 1826 atttccaatcttgcagaggaataltcaaatataaaygacagtgtactgtagaagt 1885  
Db 1590 ATTTTCCAAATTTGCCCGAGGAATATTCAAATATATATGAGACAGTACGTGAAAGATG 1649  
QY 1886 ggcacaactaatgaatccctcgtgtgcctgagcccgctatctactagtgttgtaataatg 1945  
Db 1650 GGCCAACTAATCAAGATCTCTGGGTGCTGACCGCTATTTACTGTATGTTGTAAATATGG 1709  
QY 1946 agagagatccagcttcaagactcatlgtccctctccatctgtcacaagaagatctgtag 2005  
Db 1710 AGAGAGATCTAGCTTCAGAGACTATTGGCCCTCTCCATCTGTCTACAAAGAAATCTGTAG 1769  
QY 2006 atcaaaaggaaaccagataatgtcagaacaaaggaaatgtcatcctgttctctgtatctg 2065  
Db 1770 ATCAAAAGGAAACCAAGATATGTCAAGACAAGAGAAATGTCAATCTGTTTCTGTATTTG 1829  
QY 2066 atgaagaacgaagctgtgtactcacaagagatatcaagccttctcccaatccagctg 2125  
Db 1830 ATGAGAACCCAGAGCTGTACTCTCACAGAAATATACAACTTTCTCCCAATTCAGCTG 1889  
QY 2126 gaatgcagcttgaagatcccaagatcttccaagctcccaacatcatgacaaagcatatgct 2185  
Db 1890 GAGTGCAGCTTGAAGATCCAGAGTCCAGAGCTCCAAACATCATGACAGCATCATGTGCT 1949  
QY 2186 atgtctttgtatgtctgtcagctgtgtcagttgtgtgaatgtgtgcatctgtgtatctc 2245  
Db 1950 ATGTTTGTGATGATTTGCAATGTCATGTTGTTGTGATGAGTGGGCACTACTGTACTTC 2009  
QY 2246 taagcatgtgaagcaagagagactctctctgtctctctctctctgtatataactcaaac 2305  
Db 2010 TAAAGCATTTGAGCACAGACAGTACTCTTCTTGTGCTTCTGTGATTAATCTTCAAC 2069  
QY 2306 acaaaatgtgtctatgaagaacacatcccatctccatctcagaagaaactgtgtctca 2365  
Db 2070 ACAAAATGCTTATGAAGACACACTGACCTATTCCATCTCAGAGAAACGTCTTCA 2129  
QY 2366 tgcgatgtgaaaaaaccaggtctatgtatcttgggtgtgcacaaactcgaagcttgcgaaca 2425  
Db 2130 TGTGATGTGAAAAACCCAGAGCTATGATGTTCTGGGGTCCCAACTCAGACTTTCGGNACA 2189  
QY 2426 gaggcatgtaccgctactgaaggtttctagtgtgtacaagaacactggtgatattacg 2485  
Db 2190 GAGGCAATGACCGCTTACTGAAGGTTTCTTGAATTCAGAGGGGGAGAGAGACGACG 2249  
QY 2486 aggaagatltatgaatatltca-----gcatactgtgtgtatgaanaaca 2530  
Db 2250 ACTATCTGGACCTGAGAAATATTTCACTGTAAGACAGAGCTATACATGCAATGTGTGACA 2309  
QY 2531 atgtccatgtgaacaaagaagctcttccagaatltcaagaacccctagactaggcaaaagc 2590  
Db 2310 GTCTGATTTGAACCAAGAGAGCTTCTCCAGAAATCAAGACACCTTACACTAGGCAAAAAGC 2369  
QY 2591 aatttaatgcac----- 2603  
Db 2370 AATTTAATGCCACACCAATTTCCAGAAAATGACATAGAGAAAGCTGACCTTGGTTGACAC 2429  
QY 2604 ----- 2603  
Db 2430 ACAGAACACTATAGCCTAAATATCAAAATGTCTCCTGTATGATTTGTGTATCTCTGTGC 2489  
QY 2604 ----- 2603  
Db 2490 GACAGATCTCTACTCCACATGGGCTATCCTTATCTGATCTCCAAGAAACCAATATGAGA 2549

[illegible]

|    |      |  |      |
|----|------|--|------|
| Db | 3630 | ANGTACAGTACAGGAATTGGCTGTGTTTTCACCATCTTTGATGAGACCAAAAGCTGGT       | 3689 |
| Qy | 3356 | actlccacgaaataatgaaagaanaacgcgcagggctccctgcgaatccgaatggaagatc    | 3415 |
| Db | 3690 | ACTTCACGTGAATAATATGGAAGAAAGAACTGCAGGGCTCCGTGCAATATCCAGATGGAATGC  | 3749 |
| Qy | 3416 | ccaactttaaagaataatccgtctccatgcaatgcaatggctacataatgatactaac       | 3475 |
| Db | 3750 | CCACTTTTAAAGAAATTAATTCGCTTCATCCATCAATGAGCTTACATTAATGATATACACTAC  | 3809 |
| Qy | 3476 | ctgagcttgaaatggcccaagatcaaaaggaattgatgfatccctgcctcaactgagccaca   | 3535 |
| Db | 3810 | CTGGCTTAGTAATGGCTCAGGATCAAGGAATTGATGGTATCTGCTCAGCATGGCCACGA      | 3869 |
| Qy | 3536 | atgaaacaatccattctatctatcttcaatgtagacatgcttcaatgtagcaaaataagag    | 3595 |
| Db | 3870 | ATGAAAACATCATCTATTATTCATTTCAGTGGACATGTGTTCACTGATGAGAAAAAGAGG     | 3929 |
| Qy | 3536 | agataaaatggcaatgcaatccatccatccagtgcttttgaagacatggaatgtaac        | 3655 |
| Db | 3930 | AGATTAATAATGGCACTGTACAAATCTCTATCCAGAGTGTTTTGAAGAGTGAATGATAC      | 3989 |
| Qy | 3656 | catccaagatcggaattggccgggtggaaatgacctatgggaagatccaaatgctggga      | 3715 |
| Db | 3990 | CATCCAAAGCTGGAAATTTGGCGGGGTGAATGCCCTTAATGGCAAGCATCTACATGCTGGGA   | 4049 |
| Qy | 3716 | tgagcacactttcttgctgttagacgaataatgtagaactccctcctggaaatgctcttg     | 3775 |
| Db | 4050 | TGAGCACACTTTCTTGCTGGTGTACAGCAATATGATGACACTCCCGGGAATGAGCTCTG      | 4109 |
| Qy | 3776 | gacacataagagatlttcagataacagcttcagacaataatggaatgggcccacaagc       | 3835 |
| Db | 4110 | GACACATTAGAGATTTTCAATTAACAGCTTCACGACAATATGGAAGTGGGCCCAAAAGC      | 4169 |
| Qy | 3836 | tgggcaaacctcatctatctccggatccaatgaatgctctggagacaaagagacccttctt    | 3895 |
| Db | 4170 | TGGCCACACTTCATTATTTCCGGATCAATCAATGCTCTGGAGCAACCAAGAGCCCTTTTCTT   | 4299 |
| Qy | 3896 | ggatcaagtggaatcgtgttgacaccaaataatltcaagagatcaagaaccacaggtgtcc    | 3955 |
| Db | 4230 | GGATCAAGGTGATCTGTGTGGACCAATGATTAATTCACGGCATCAAGACCAGGGTGCC       | 4289 |
| Qy | 3956 | gtcgaagatltccagcctcaactcatctctcagttatctatctatgatatgctcttgatga    | 4015 |
| Db | 4290 | GTCGAAGTTCTCCAGCCTCTACACTCTCTCAGTTTATCAATCATGATATGCTCTGATGGGA    | 4349 |
| Qy | 4016 | agaagtggcaacttaccggaagaattccaactggaaccttaatgctctctcttgcaatg      | 4075 |
| Db | 4350 | AGAAATGGCAGACTTATCGAGGAATTTCCACTGGAACTTAATGGTCTTTTGGCAATG        | 4409 |
| Qy | 4076 | tgatltcatctggatataaaacacaaatattttaaaccctcaatlatgtctcgataatccc    | 4135 |
| Db | 4410 | TGATTTCATCTGGGATTAACACAAATATTTTAACTCCATTAATATGCTGATACATCC        | 4469 |
| Qy | 4136 | gtttgcacccaactcatataatgaatcgcgaacctcttcgaatgagttgatgggctgtg      | 4195 |
| Db | 4470 | GTTTGCACCCCAACTCAATTAATGAGATTTCCGAGCACTTTTGCTGATGAGTTGATGGGCTGTG | 4529 |
| Qy | 4196 | atttaaatagttgacgaatgcattgggaatggagatgaagaatataatcaatgacga        | 4255 |
| Db | 4530 | ATTTAATAGTTGCAGATGCCATTGGGAATGGAGATGAACCAATATTCAGATGCACAGA       | 4589 |
| Qy | 4256 | ttaactgttcaatccatctattacaataatgtttgcacatgctgtctccctcaaaagctcgac  | 4315 |
| Db | 4590 | TTACTGTCTATCTCTACTTATACCAATATGTTTGGCCACTGCTCTCTTCAAAAAGCTGCAC    | 4649 |
| Qy | 4316 | ttaacatcccaaggagagatgaatgctctggaagacatccagttgaataatccaaagatgggc  | 4375 |
| Db | 4650 | TTTCACTTCACAAAGGAGAGTAATGCTTGAGAGACTCAGGTGATTAATCAAAAAGTGGGC     | 4709 |
| Qy | 4376 | tgcgaatggagcttccaaataaacaatgaagaatcaagatgaactactcaaggagtaaat     | 4435 |

Db 4710 TCGAAGTGGACTTCAGAGACAAATGAAAGTCACAGGAGTAACTACTCAGGAGTAAAT 4769  
QY 4436 ctctgctaccagatgtatgtgaaggaagttccctcatctccagcagcagatggccatc 4495  
Db 4770 CTCTGCTTACAGATGATGTGAAGAGTTCCTCATCTCCAGACAGCAAGATGGCATC 4429  
QY 4496 agtggatctctcttttcagaatgcaaaatgaagtttttcaggaggaatcaagactcct 4555  
Db 4830 AGTGACCTCTTTTTCAGATGCAAAAGTATTAAGTTCAGGGAATCAAGACTCCT 4489  
QY 4556 tcaacctgtgtgaactctcttagaccaccgttactgactgctactctgaatcaacc 4615  
Db 4890 TCACACCTGTGGTGAACCTCTTAACCCACCGTACTGACTCGCTACTTCGAATTACCC 4949  
QY 4616 cccagagttgtgaccagaattgcccctgaagatgaggttctgggtctgcaaggaacag 4675  
Db 4950 CCCGAGTGGGTGACACAGATGGCCCTGAGAGTGGAGTGTGGGCTGAGGACACAGG 5009  
QY 4676 acctactactgacagcagatc 4699  
Db 5010 ACCTCTACTGAGGTGGCCACTGC 5033

## RESULT 10

US-09-243-539-1  
; Sequence 1, Application US/09243539  
; Patent No. 6130203  
; GENERAL INFORMATION:  
; APPLICANT: VOORBERG, Johannes J.  
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/243,539  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/558,107  
; FILING DATE: 13-NOV-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: ISACSON, John P.  
; REGISTRATION NUMBER: 33,715  
; REFERENCE/DOCKET NUMBER: 30472/212  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ. ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5035 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 35..5017  
US-09-243-539-1

Query Match 78.7%; Score 3748; DB 3; Length 5035;  
Best Local Similarity 87.4%; Pred. No. 0;  
Matches 4374; Conservative 0; Mismatches 60; Indels 570; Gaps 2;

QY 266 ccacacatgcaatatagagctctccaccgtctctctctgctcttgagctttctgttta 325  
Db 30 CCACATGGAATATGACCTCTCCACCTGCTTTCTGTGCTTTGGCGATTTGCTTTA 89  
QY 326 gtgcacacaaataatactactggtgtgacgtggaactgctatctggaactatgcaagtc 385  
Db 90 GTGCACCAAGATVCTACTGGGTGCAATGGAACTGTCATGGGACTATATGCCAAAGTG 149  
QY 386 atctcggtgagctgtcctgtggaacgaagatctctctctagaagtgcacaaatctttccat 445  
Db 150 ATCTGGGTGAGCTGCTGTGTGAGCGCAAGATTCTCTCTTAAGTGCACAAATCTTTTCAT 209  
QY 446 tcaaacacctgaactcgtgttaaaaaaacctgttttttagaattcgaagttccacttttca 505  
Db 210 TCACACCTCTAGCTGCTGTACAAAAGACTGTGTTGTAGATTCCAGGATCACCTTTTCA 289  
QY 506 acatcgctaaagccaagccaccctgtgagtgtctgtaagttccatcacatccaggctgaag 565  
Db 270 ACATCGGTAAAGCCAAAGCCACCTGGATGGCTGTGCTGATGCTTACCATCCAGCTGAGG 329  
QY 566 ttatgatatacagtgtgtcaattacaattagaacaatggtctcccatcctgtcagttcaatg 625  
Db 330 TTTATGATACAGTGTGATACACTTAAGAAACATGAGCTTCCATCCCTGTCAGCTTCATG 389  
QY 626 ctgtgtgtgtatctctactggaagaactctctgagggagctgaatatgatatcaagaactc 685  
Db 390 CTGTGGGTGTATCTCTACTGGAAGCTTCTGAGGAGCTGAATGTGATGATCAACACAGTC 449  
QY 686 aagggagaagaagaatgtataaagctctccctgtgtgaacccaataatgctgtgcag 745  
Db 450 AAAGGGAAGAAGATGATGAATGAATGCTCCCTGTGTGAACCATACATATGCTGGCAGG 509  
QY 746 tccatgaagaagaatggtlccaaatggtcctctgaaccactgtgcttactactatcttt 805  
Db 510 TCCGTGAAGAGATGTGTCATGAGGAGCTGTGACCCACTGTCCTTACTACTCATACTTTT 569  
QY 806 ctatctggaacctggttaaaagacttgaattcaagcctctcaattggaacctactgtatgta 865  
Db 570 CTCTATGGACCTGTGTAAAGACTTGAATTCAGGCTCATTTGAGGCTTACTATGATGTA 629  
QY 866 gaagaaggagctctgccaaggaagaacacagacctctgcaaatlatactactttt 925  
Db 630 GAGAAGGAGACTTGGCCAAAGGAACACAGACCTTGGCAAAATTTATVACTCTTTTG 689  
QY 926 ctgtatttgaagaaggaaagtgtgcaactcagaacaaagaactctgtatgacagata 985  
Db 690 CTGTATTTGATGAAGGAAAGTTGGCAGTCAAGAAACAAAGAACTCTTGATGCAGATA 749  
QY 986 gggatgctgcatctgtctgggacctgacctaaatgcaacagcaatggttatgttaaca 1045  
Db 750 GGGATGCTGCATCTGCTCGGGCCCTGGCTTAAATGACACAGTCAATGTTATGTAAACA 809  
QY 1046 ggtcctgccaagctctgatatgtgacaggaatcagtcatacttgcatgtattgtaa 1105  
Db 810 GGTCTCTGCCAGGTCTGATTTGATGGCACGAAATCAGTCTTTGGCATGTATTGGAA 869  
QY 1106 tgggcaaccactctctgaagtgcataatctctctgaaggtcacaacattctgtgaa 1165  
Db 870 TGGGCACCACTCTGAAGTGCACATCAATATTCCTCAAGGTGCACACATTTCTGTGAGA 929  
QY 1166 accatgccaaggtctctctggaatctgcgaataactttctctactgtcacaacactc 1225  
Db 930 ACCATGCCACGGCTCTTGGAAATCTCGCAATVAACTTCTTACTGCTCAACACTCT 989  
QY 1226 tgaatgaccttgacagattctactgttttgtcatactctctccaccaacatgtagca 1285  
Db 990 TGATGACCTTGGACAGTTTCTACTGTTTGTGATATCTCTTCCACCAATGATGAGCA 1049  
QY 1286 tgaagactaatgtcaaaagttagacagctgtccagaaggaaacccaactagaatgaataa 1345  
Db 1050 TGGAACTTATGTCAAAAGTAGACAGCTGTCCAGAGGAACCCCAATCAATGAATGAATAA 1109

[illegible]

|    |      |  |      |
|----|------|--|------|
| Db | 2130 | GAGGCATGACCGCCTTACCTGAGGTTCTAGTTGTATTCACAGGGGGGAGGAGCGACG      | 2249 |
| Qy | 2486 | agaaacgctctatgaagataattca-----gaactctgcagtaaaaca               | 2530 |
| Db | 2250 | ACTATCTGAGACCTGGAGAAGATATTGCTGTAAGACGACGATCTACGACATCGTCGACA    | 2309 |
| Qy | 2531 | atgcattgaacccaagaagctctccagaattcaagaacccctagactaggcaaaagc      | 2590 |
| Db | 2310 | GTCGTGATTGAACCAAGAAGCTTCTCCGAAATTCAAGACACCTTAGCTAGGCAAAAAGC    | 2369 |
| Qy | 2591 | aatttaagccac-----  | 2603 |
| Db | 2370 | AATTTAATGCCACCACAATTCCAGAAAATGACATAGAGAAGACTGACCTTGTTGCAC      | 2429 |
| Qy | 2604 | -----  | 2603 |
| Db | 2430 | ACAGACACACTATGCTTAATACAAAATGTCCTCTAGTATTGTTGATGCTCTTGC         | 2489 |
| Qy | 2604 | -----  | 2603 |
| Db | 2490 | GACAGAGCTCTACTCCACATGGGCTATCTTATCTGATCTCCAGAAAGCAAAATATGAGA    | 2549 |
| Qy | 2604 | -----  | 2603 |
| Db | 2550 | CTTTTCTGATGATTCATCACCTGGAGCAATAGACAGTAATAACAGCCTGTCTGAATGA     | 2609 |
| Qy | 2604 | -----  | 2603 |
| Db | 2610 | CACACCTTCAGGCCACACCTCCATCCAGTGGGAGACATGGATTATTAACCCCTGATCAAGCC | 2669 |
| Qy | 2604 | -----  | 2603 |
| Db | 2670 | TCCAAATTAAGATTAATAGAGAAACTGGGGACAACCTGACATCCTCTTGCTGGATTAAC    | 2729 |
| Qy | 2604 | -----  | 2603 |
| Db | 2730 | ACTATGCTACTCAGATACCAAAAGAGAGTGAATCCAGAGAGAACTCACCAAAAAA        | 2789 |
| Qy | 2604 | -----  | 2603 |
| Db | 2790 | CAGCTTTAAGAAAAAGAGATTCATTTTTGCTCCAGAACCTTGTTGAACCAATTCATGCA    | 2849 |
| Qy | 2604 | -----  | 2603 |
| Db | 2850 | TAGCAGCATTAATGAGGAGCAAAATTAAGCCGAATAGAACTGCTGGCAAAAGCAAG       | 2909 |
| Qy | 2604 | -----cccaacagcttgaaagccaaacaaagggaa                            | 2635 |
| Db | 2910 | GTAGACTGAAGGCTGTGCTCTCCAAAACCCACCACTCTTGAAGCCCATCAACGGGAA      | 2968 |
| Qy | 2636 | taactcgtactactcctcagtcagatcaagaagaaattgaactatgatagaataacatacag | 2695 |
| Db | 2970 | TAACTCGTACTACTCTTCAGTCAGATCAAGAAAGAAATGACATATGATGATNCATATACG   | 3029 |
| Qy | 2636 | ctgaatgaagaagaagattctgaactatgatagaatgaagaatgaagaatgaagcccgca   | 2755 |
| Db | 3030 | TTGAAATGAGAAAGAGAGATTTTGACATTATGATGAGATGAATAATCGAGACCCGCCCA    | 3089 |
| Qy | 2756 | gcttcaaaagaaacagacactatttattgtctcagatggagagagccttggaattatg     | 2815 |
| Db | 3090 | GCTTTCAAAAGAAACACGACACTATTTTATTTCTCCAGTGGAGAGGCTCTGGAAATTATG   | 3149 |
| Qy | 2816 | ggaatgaatagctccccaactgctcctaagaacagggctcagaatgagcagtgccctcaag  | 2875 |
| Db | 3150 | GGATGAGTAGCTCCCCACATGTTCTTAAGAAACAGGGCTCAGAGTGGCAATGTCCCTCAGT  | 3209 |
| Qy | 2876 | tcaagaagaatgctttccaggaattactgaatgagctcccttaactcaagccctatacagtg | 2935 |
| Db | 3210 | TCAAGAAAGTTGTTTCCAGGAATTAATGATGGCTCTTTACTCAGGCCCTTAATACCGTG    | 3269 |
| Qy | 2936 | gagaactaaatgaacatttggagctccctgggccaataataagacagaagttaagaata    | 2995 |

Db 3270 GAGAACTAAATGAACTTTGGGACCTCGGGCCATATATAGAGCAGAAATTGAAGATA 3329  
Oy 2296 atatcatgtaacttcaagaatcaggcctctgctccctatctctctatctagcctta 3055  
Db 3330 ATATCATGTAAGTCTTGAAGATGAGGCTCTGCTCCATATCTTCTATCTTGAAGCTTA 3389  
Oy 3056 ttctctatggaagatcagaagcagaagcagaacctaagaaacttctcaagccta 3115  
Db 3390 TTTCTTATGAGGAAGATCAGAGCAGAGACCACTAAGAAAACTTTGTCAAGCTTA 3449  
Oy 3116 atgaacccaacttacttcttggaagtgcaacatactatggaacccaactaagatgagt 3175  
Db 3450 ATGAACCAAACTTACTTTTGGAAATGCAACATCATATGACACCCACTAAAGATAGT 3509  
Oy 3176 ttgacgtcaagcctggccttattctctgagtgttgaccctggaaagaaatgctgaccag 3235  
Db 3510 TTGACGTCAAGCCTGGGCTTATTTCTGTATGTGACCTGGAAAAAGATGTGCACACAG 3569  
Oy 3236 gctgtaattgacccctctgctctgcaactaacaacactgaaacctgctcatggaagac 3295  
Db 3570 GCCTGATGGAACCCCTTCTGTCTGCTGCACACTAACACACTGMAOCCCTGCTCATGGAGAC 3629  
Oy 3296 aagtgacagtaagaaattgctctgcttcttccacatcttgatgagacccaagctggt 3355  
Db 3630 AAGTGACAGTACAGGAATTTGCTGTGTTTTCACATCTTTGATGAGACCAAAAGCTGCT 3689  
Oy 3356 acttccctgaaataatggaagaagaaactgcaagctccctgcaataaccagaatggaagatc 3415  
Db 3690 ACTTCACGTGAAATATGGAAGAAAGAACTGCAAGGCTCCCTCAATATCCAAATGGAATATC 3749  
Oy 3416 ccaactttaaagaagaattatcgcttccatgcaatcaatgtaactaataatgatacactac 3475  
Db 3750 CCACCTTTTAAAGAGANTATACGCTTCCATGCAATCAATGCTCATATATGATACACTAC 3809  
Oy 3476 ctgagcttagtaatgctcagaatcaagaatctgatgtaactctgctcagaatgagca 3535  
Db 3810 CTGCGTAGTAATAGCTCAGATCAAAAGATGATGATGCTGCTCAGCATGGGAGCA 3869  
Oy 3536 atgaacacatcattatctatcttcaagtggagatgcttcaactgtaagaaaaagag 3595  
Db 3870 ATGAAACATCCTATCTATTCATTCAGTGAGCATGTGCTACTGTACGAAAAAAGAG 3629  
Oy 3596 agataaaatgacatgtaacatctatccatccagtgcttcttgagacagtgaagatgtac 3655  
Db 3930 AGTATATAATGAGCAGCTGTACAAATCTCTATCCAGGTGTTTTGAGACAGTGAATGTAC 3689  
Oy 3656 catcaaaagctggaatttggcgggtggaatgctctatcttgagacatctacatgctgga 3715  
Db 3990 CATCCAAAGCTGGAATTTGGCGGGTGAATGCTTATTTGGCAGCATCTCATGCTGGGA 4049  
Oy 3716 tgagcacacttctctggtgtgtacagaataagtgtcaagactccctgggaatgctctg 3775  
Db 4050 TGAGCACACTTTTGTGTGTACAGCAATAGTGTAGACTCCCTGGGAATGGCTTCTG 4109  
Oy 3776 gacacattagaatttcaagattacagctcagaacaataatgagaagtgagcccaag 3835  
Db 4110 GAACATTAAGATTTCAGATTACAGCTTCAGACAAATATGACAGTGGGCCCAAGC 4169  
Oy 3836 tggcgaactcattatccggaatcaatcaatgcttggagacccaaggaaccccttctt 3895  
Db 4170 TGGCCAGACTTCTATTTCCGGATCAATCAATGCTGAGACCAAGAGACCCCTTTCTT 4229  
Oy 3896 ggaatcaagtgatctggttggaaccaatgattatcaagcacaagacccaagtgccc 3955  
Db 4230 GGATCAAGGGGATGCTGTGGACCAATGATTTTACAGGCAATCAAGACCAAGGTCGCC 4289  
Oy 3956 gtaagaagtctccagcctcatatcctcagttatcaacatgataagcttgatgagga 4015  
Db 4290 GTCAAGAGTCTCTCAGGCTCTACATCTCTAGTTATCATCATGTATAGCTTGATGGGA 4049  
Oy 4016 agaagtggaacttatacgaagaatccacatggaacctaagtgcttcttggaagt 4075  
Db 4350 AGAAGTGGCAGACTTATCGAGAAATTCACCTGAACCTTAATGATGCTTCTTGGCAATG 4409

Oy 4076 tggatcatctgggaataaacaataatttaacccctccaattatgtcgtatcatcc 4135  
Db 4410 TGGATTCATCTGGGATTAACACAAATATTTTAACTCCATTAATATGCTGATACATCC 4469  
Oy 4136 gtttgcacccaactcatalatagatcaltgcagcactcttcgcagtggaatgatagtgctg 4195  
Db 4470 GTTTGCACCCCAACTTATATAGCATTTGCGAGCACTCTTCCATATGAGATTTATGGGCTG 4529  
Oy 4196 atttaaatgttgcaagatgccaattgggaatggaagtaagaataatcagatgacaga 4255  
Db 4530 ATTAAATATGTTGAGCATTCATTTGGGAATGAGATGAAGCAATATCATGATGCACAGA 4589  
Oy 4256 ttactgcttcatctacttcaacaataatgtttgccaactgctctctcaaaagctgac 4315  
Db 4590 TTACTGCTTATCTCTATCTTAAATATATGTTTGGCACTGCTGCTCTTCAAAAGCTGAC 4649  
Oy 4316 ttcaacctcaaggagaagatgtaatgcttgagacactcagtgtaataatccaaagatgagc 4375  
Db 4650 TTCACTCCCAAGGAGAGATATGCTTGAGACTCTGAGTGAATATCCAAAGAGTGGC 4709  
Oy 4376 tgaagtgaacttccagaagaatgaaagtcaagaagtaactactcaaggaagtaaat 4435  
Db 4710 TGCAAGTGAAGTTCACAGAACATGAAGTCAAGAGTAACTACAGGAGTAAAT 4769  
Oy 4436 ctctgcttaccagatgatagtgaagagttctctcatctccagagtcgaagtgcacatc 4495  
Db 4770 CTCTGCTTACAGATGATATGTAAGAGTTCCTCACTCTCAGCAAGATGGCCATC 4829  
Oy 4496 agtggacatcctcttctcaagaatggaagaaagtcttccagggaaatccaagactcct 4555  
Db 4830 AGTGAATCTCTTTTTCAGAAATGGAAGTAAAGTTTTCAGGGAATCAACAGACTCT 4889  
Oy 4556 tcaaacctgtgtggaactctctagaccaccagtaactgtaactgcttcaactcacc 4615  
Db 4890 TCACACCTGTGGTGAAGTCTCTAGACCCAGCTTACTGACTCGCTTCAATTCACC 4949  
Oy 4616 cccaagatgggtggaacaaattgcccctgaggaatggaaggtcttgggttggaagcaag 4675  
Db 4950 CCCAGATTTGGGTGACCAATATGCTCTGAGATGAGGTCTGGGCTGAGGACACAGG 5009  
Oy 4676 acctctactgactcagagagatc 4699  
Db 5010 ACCTTACTAGGGGTGGCCACTGC 5033

RESULT 11  
US-08-670-707A-38  
; Sequence 38, Application US/0867070A  
; Patent No. 5859204  
; GENERAL INFORMATION:  
; APPLICANT: Lollar, John S.  
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Circle Suite 201  
; CITY: Boulder  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/670,707A  
; FILING DATE: 26-JUN-1996  
; CLASSIFICATION: 435  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/US94/13200  
; FILING DATE: 15-NOV-1994





|    |      |  |      |
|----|------|--|------|
| Db | 1660 | AAAGATCTGGCTTGGGACACTTGGCCCTCTCTCATCTGCTTACAAAGAACTGTAGAC          | 1739 |
| Qy | 2008 | caaaaggaacccagataaactgttcagaacaagagaaatgtcatctctgttctgtatctgat     | 2067 |
| Db | 1740 | CAAAAGAGAAACCAAGATGATGTCTAGACAAGAAAGCAAGCATCTCTTTCTGTATTCGAT       | 1799 |
| Qy | 2068 | gagaaacgaagctgtgtactcctcacagaaatatacaacgccttctcccaatccacgttga      | 2127 |
| Db | 1800 | GAGATCAAAAGCTGTGTAACCTCGCAGAGAAATATTCAGCGCTTCTCCCAATCCGGATGGA      | 1859 |
| Qy | 2128 | gtgagacttgagatccagagttccaaagcttccaaatcatatgtacaaagatcaatgtgcctat   | 2187 |
| Db | 1860 | TTACAGCGCCCAAGGATCCAGAGTTCACAGCTTTACATCATGACACAGATCAATAGGCTAT      | 1919 |
| Qy | 2188 | gttttcatagtttgcaggtgttcagtttctgttcgaatgaatgtacgttgcatacttcta       | 2247 |
| Db | 1920 | GTTTTCATACCTTCGACGCTGTGGGTTTGTTCGACAGAGGTGCACTATGTTACATTCTA        | 1979 |
| Qy | 2248 | agcatvgagacaagactgtactccttctgtctcttctctctgtatacttccataaac          | 2307 |
| Db | 1980 | AGTGTGGAGACAGACGAGACTTCCTCCTCGCTTCTTCTGCTTACACCTTCAAACAC           | 2039 |
| Qy | 2308 | aaaatgtctataagaacacacatccactatcccatctcgaaggaaactgtcttcattg         | 2367 |
| Db | 2040 | AAAATGCTCATGAAAGCACACACACACCCCTTCCCTCTCAGGAGAAACGGTCTTCATG         | 2099 |
| Qy | 2368 | tcgactggaaaacccaagttcctatgtatctctcgtgggtggccaaactgaagcttccgaaaga   | 2427 |
| Db | 2100 | TCAATGGAANAACCAAGGCTCTGTGGTCTCTAGGGTGCACAACTCAGACTTCGGAAACGA       | 2159 |
| Qy | 2428 | ggcactgacgccttactgaaggttcttgaatcttgagaacaacactggtactatctacag       | 2487 |
| Db | 2160 | GGGATGACACCTTACTGAAAGGTATTAATTGTGACAGGGAATGTGATATTATTAAGAC         | 2219 |
| Qy | 2488 | gacacgttatgaagatatttcaagcatacttgcgtgagtaaaacaatgcatctgaaccaga      | 2547 |
| Db | 2220 | AACACTTATGAAATATTTCCAGGCTTCTTGCTGAGTGGAAAGAAATGTCATTGAAACCGAGA     | 2279 |
| Qy | 2548 | agctctcccaagatccaagacaacctagcacttagcaaaagcaattaatgtccacccca        | 2607 |
| Db | 2280 | -----  | 2279 |
| Qy | 2608 | ccagttcttgaacgcctcaacacggagaaatactcgttactacttctcgaatccagaag        | 2667 |
| Db | 2280 | -----GACATTAACGCTTCCACTTTCAACCGGAGAGAAAC                           | 2315 |
| Qy | 2668 | gaaatgtactatgtatataccacataatcagttgaatgaagaagaagatlttgacattat       | 2727 |
| Db | 2316 | AAAATGACATATGATATATCTTCTCMACTGAAACGAAGGAGAGATTTTGACATTTTAC         | 2375 |
| Qy | 2728 | gattagagataaaatacgaagccccgcgcacatttcaaaagaaacacgcacattattat        | 2787 |
| Db | 2376 | GGTATGAGTATAAATATCAGACCCCTCCACACTTTTCGAAGAGAACCCGACACTTATTTCAAT    | 2435 |
| Qy | 2788 | gcttcagatvgagagagctctcbbggaattatggtgatgtatagcttccccaactgtcttaagaac | 2847 |
| Db | 2436 | GCTTCGGTGGAGCAAGCTCTGGGATTTACGGGATGTAGACGAATCCCCCGGGCCCTTAAGAAC    | 2495 |
| Qy | 2848 | aggagctcagaagtgcgcagtgctccctcagttcagaagaagtgttttccagaaattactgat    | 2907 |
| Db | 2496 | AGGGCTCAGAACGAGAGAGGCTCTCGATTCAAGAAAGGTGTCTTCGGGAAATTTGCTGAC       | 2555 |
| Qy | 2908 | ggtcctcttctacagcccttatacgittgagaaactaaatgaacatttggagctctcctggg     | 2967 |
| Db | 2556 | GGGCTCTTTCACGAGCGCTGTGACCGCGGGGAATCAACAAACATTCGGGCTTTGGGA          | 2615 |
| Qy | 2968 | ccatataaagaacagaagtittgaagaataatcatcgttgaactttcagaatccagagctct     | 3027 |
| Db | 2616 | CCCTACATCAGACGGGAAGTGAAGACACATCATGTGTACTTTCAAAAACCAAGCGCTT         | 2675 |
| Qy | 3028 | cgctccattctctatttctagccttatttcttatatgagaagatccagaggaagagaca        | 3087 |

|    |      |   |      |
|----|------|---|------|
| Db | 2676 | CGTCCCTATTTCCTTCTACTGAGCCCTTAATTCTTATCCGGATGATCAGGACGACGGGCA        | 2735 |
| Qy | 3088 | gaacctgaaaaaaaccttctgcaagccccaatgaaacccaacacttacttcttgaaagtgcga     | 3147 |
| Db | 2736 | GAACCTGCACACACTTCGTGTACAGCCAAATGAANACGAACTTACTTTTGGAAACTGCAG        | 2795 |
| Qy | 3148 | catcaatgycacccactaaagatgagtttgcactgcaaaagcctggacttattctctgcat       | 3207 |
| Db | 2796 | CATCACAATGGCACCCACAGAAAGACGATTTGACTGCAGAAAGCCCTGGCCCTACTTTCTGAT     | 2855 |
| Qy | 3208 | gttgaacctggaaaaagtgtggaactgaagcgttatttggccccctctgtctgcgaacact       | 3267 |
| Db | 2856 | GTTCACCTGGAAAAAGATGTGCACTCAGGCTTATGTATGGCCCTTCTGATCTGCGCGCC         | 2915 |
| Qy | 3268 | aacacactgaacctgtctctatgvgagaacaagtacacagatgcagaatttgcctgtcttc       | 3327 |
| Db | 2916 | AACACCTCGAACCGCTGCTCAACGGTAGACAGTAGTACCCTGTCAAGATTTGCTGTGTTTTTC     | 2975 |
| Qy | 3328 | accacctcttgatgagaccaaagaactgttacttcaactgaaataatgynaagaactcagag      | 3387 |
| Db | 2976 | ACTATTTTTGTATGAGACAAAGAGCTGTACTTCTACTGAAAAATGTGGAAAGAACTGCGG        | 3035 |
| Qy | 3388 | gtctccctgcaatatcccaatgvgaaagatcccaactttaaagaatatactgcctcctatgca     | 3447 |
| Db | 3036 | GCCCCCTGCCACCTGCGAGATGTGAAGACCCACACTCAAAAGAAAACATATGCTTCCATCA       | 3095 |
| Qy | 3448 | atcaatgtcctcaataatgvgatacactccctgtgttagtaatgtgtccaagatacaagaagt     | 3507 |
| Db | 3096 | ATCAATGGCTTATGTGATGATGATACACTTCCTGGCTTAGTATGGCTCAGATATAAAGATC       | 3155 |
| Qy | 3508 | cgatgltatctcgtctcagcatgvggcagcaatgynaacatccatcttcaatcttcaatgtaga    | 3567 |
| Db | 3156 | CGATGTATCTGCTGCACATGGGGAGCAATGAAAATATCATTCGATTTCATTTTAGCGGA         | 3215 |
| Qy | 3568 | catgtgtcactgtgacgaaaaaagaagagatataaaatgycactgacaactctctatcca        | 3627 |
| Db | 3216 | CACGCTTCACCTGTACGAAAAAGAGAGATATTAATAATGGCGGTACACTTCATATCCG          | 3275 |
| Qy | 3628 | ggtgtcttttttgagaagatgynaatgltataccatccaagaagcgggaatttggcgggtggaatgc | 3687 |
| Db | 3276 | GGTGTCTTTGAGACAGTGGAAATGTCTACGCTCCAAAGTTTGGAAATTTGGCGAATAGAAATGC    | 3335 |
| Qy | 3688 | cttatctggcagcatctactatgctgvgatgagacacttcttctgtgtlatacgaacaatag      | 3747 |
| Db | 3336 | CTGATTTGGCGAGCACCTGCAAGCGTGGGATGAGCAGCAGCTTTCCTGTGTACAGCAAGAG       | 3395 |
| Qy | 3748 | tgtaagactccccctgggaatgagctctctgacacattagagaatttcagaattacagcttca     | 3807 |
| Db | 3396 | TGTACAGGCTCCACTCTGGGAATGTGGCTTCTGGACGATTAGAGATTTTTCAGATCACACTTCA    | 3455 |
| Qy | 3808 | ggacaatatggaagatgvggcccccaagaactbvgccaagactcatattcccgatcaactaat     | 3867 |
| Db | 3456 | GGACAGTAGTGACAGTGGGGCCCAAAAGCTGGCCAGACTTATTAATTCGGATCAATTCAT        | 3515 |
| Qy | 3868 | gcccggagcaacaagvggcccccttctcttgatcaaaagtgtgactctgttgagcaacaatgatt   | 3927 |
| Db | 3516 | GCCGTGAGACACACAGATGCCCACTCCTGTGATTAAGTGTGATCTGTGGCACACATGATGC       | 3575 |
| Qy | 3928 | atccaagcatcaagacccaagggtgcocgtcagaagttctccagcctctacactcctcag        | 3987 |
| Db | 3576 | ATTTACGGCATCATGACCCAGGGGTGCCCGCTGCAAAATTTTTCAGCCTCTACATCTCCAG       | 3635 |
| Qy | 3988 | tttctaactatgatagtactgttgatgggaagaagtggcagacttaccggaagaatttccact     | 4047 |
| Db | 3636 | TTTATTCATCATGTACACTCTTTGACGGAGAGAACTGGCAAGTTTACCGAGGGCAATTCACG      | 3695 |
| Qy | 4048 | ggaaccttaaatggtcctctcttgacaatgvtgattactcggagataaacaataattttt        | 4107 |
| Db | 3696 | GGCACCTTAATAGGTCTTCTTTGGCAATGTGAGACGATCTGGGATTAACACATATTTTTT        | 3755 |
| Qy | 4108 | aacctccaattatgtctgatacatccgttttgaccaaccaactcattatagatctgcagc        | 4167 |
| Db | 3756 | AACCTTCGATTTGGTGGCTGCGTACATCCGTTTGGACCCCAACATTAACGATCCGCGAGC        | 3815 |



[illegible]

RESULT 12  
US-09-037-601-38  
Sequence 38, Application US/09037601  
Patent No. 6180371  
GENERAL INFORMATION:  
APPLICANT: Iollar, John S.  
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Circle Suite 201  
CITY: Boulder  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/037,601  
FILING DATE: 26-JUN-1996  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US94/13200  
FILING DATE: 15-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/212,133  
FILING DATE: 11-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/864,004  
FILING DATE: 07-APR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferber, Donna M.

```

: REGISTRATION NUMBER: 33,878
: REFERENCE/DOCKET NUMBER: 75-95F
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 303/499-8080
: TELEFAX: 303/499-8089
: INFORMATION FOR SEQ ID NO: 38:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4334 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: not relevant
: MOLECULE TYPE: cDNA to mRNA
: HYPOTHETICAL: NO
: ORIGINAL SOURCE:
: INDIVIDUAL ISOLATE: Factor VIII lacking B domain
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 3..4334
:
: US-09-037-601-38

```

[illegible]

[illegible]

|    |      |   |          |              |                                 |      |
|----|------|---|----------|--------------|---------------------------------|------|
| Db | 1800 | GGAATTC   | AAAGCTGG | TACCTCGCAGAA | AAATTTACAGCGCTTCTCCCAATCCGATGGA | 1859 |
| Qy | 2128 | gtcgagcttggagatccagagttccagagctccacaatcagacagcaatcgaatgcat        |          |              |                                 | 2187 |
| Db | 1860 | TTACAGCCCCAGGATCCAGATTCGAAAGCTTTACATCATGACAGCATCAATGGCTAT         |          |              |                                 | 1919 |
| Qy | 2188 | gtttttgtagtcttcgagttgtcagttgtgtttgcatgagtgycatactgtacata          |          |              |                                 | 2247 |
| Db | 1920 | GTTTTGTATACCTTTGCAGAGCTGTGGTTGTTCACAGAGGTGGCATCTGTATCTCTA         |          |              |                                 | 1979 |
| Qy | 2248 | agcatttggacagcagactgaactcttccttctgtcttcttctctcgtgatataccttcaaac   |          |              |                                 | 2307 |
| Db | 1980 | AGTGTGGAGACAGACGCGAGCTTCCTCCGTCTCTCTCTGCGTACACCTTTCAAAAC          |          |              |                                 | 2039 |
| Qy | 2308 | aaaatggtccttggagaacacatcaccttatcccatcttcagtgagaacatcttctcatg      |          |              |                                 | 2367 |
| Db | 2040 | AAAATGGCTTATGGAAGACACACTACCTGTTCCTCCCTTCTCAGAGAAACGGTCTTCATG      |          |              |                                 | 2099 |
| Qy | 2368 | tcgatggaataaacccagagttctatgtagttcttggtgtgcacaactcagacttccgaacaga  |          |              |                                 | 2427 |
| Db | 2100 | TCAATGGAAACCCAGGTCCTGTGGCTCTTGGGTGGCCACAACACTTGGGGAACAGA          |          |              |                                 | 2159 |
| Qy | 2428 | ggcatgaccgccttacttgaagagtttctagtgtgtgacaagaacactggtgtatattacgag   |          |              |                                 | 2487 |
| Db | 2160 | GGGATGACAGCCTTACTGTAAGGTGTATAGTTGTGACAGGACATTTGCTATTATATGAC       |          |              |                                 | 2219 |
| Qy | 2488 | gacagtttggagaatattcttgcagacttctgtatgtatgaataaaatccattgaaccaaga    |          |              |                                 | 2547 |
| Db | 2220 | AACACTTATGAAAGATATTCCAGGCTTTCTCTAGTGGGAAGAGATGTCAATTAAACCCAGA     |          |              |                                 | 2279 |
| Qy | 2548 | agctcttcccgaaatltcaagaacacctagcactaggcaaaagcaatltatgtccaccca      |          |              |                                 | 2607 |
| Db | 2280 | -----   | -----    | -----        | -----                           | 2279 |
| Qy | 2608 | ccagtcttgaagcgcacatcaacggygaataactcgttactacttccagtccagatacgaag    |          |              |                                 | 2667 |
| Db | 2280 | -----   | -----    | -----        | -----                           | 2315 |
| Qy | 2668 | gaatttgcatatgtataccataatcaagtgtgaatgaagaagaagatttgcattat          |          |              |                                 | 2727 |
| Db | 2316 | AAAATGGACTTGTGATATCTTCTCACTGGAACAGAGGAGAAATTTTACATTTTAC           |          |              |                                 | 2375 |
| Qy | 2728 | galtgagatcgaaatltcagagcccccgcgagtttccaagaagaacaacgacacttattat     |          |              |                                 | 2787 |
| Db | 2376 | GGTAGAGATGAAATTCAGAGACCTCGAGCTTTCAGAAAGAGAAACCGAACAATTTCAT        |          |              |                                 | 2435 |
| Qy | 2788 | gctgcagtgagagagctctggtgattatlyggaatgtagtactccccaatlytltctaaagaa   |          |              |                                 | 2847 |
| Db | 2436 | GCTGGGGTGGAGAGCTCTGGGATATTACGGGATGAGCAATCCCCCGGGCGCTTAAGAAAC      |          |              |                                 | 2485 |
| Qy | 2848 | agggtcgaagtagtgacatgtgtccctcaagttcaagaagaatgtgttttccaggaattactgat |          |              |                                 | 2907 |
| Db | 2496 | AGGGCTTCGAACGGAGAGGTGCTCGGTTTCMAAAGTGTGCTTCCGGAAATTTCTAC          |          |              |                                 | 2555 |
| Qy | 2908 | ggtcccttactcagcccttataacggtlygagaactaaatgaaacatttggagactccctggg   |          |              |                                 | 2967 |
| Db | 2556 | GGCTCTTTCACAGCGCGGTGTACCGGGGGAACCTCAAAACACTTGGGCGCTCTTGGGA        |          |              |                                 | 2615 |
| Qy | 2968 | ccatatataagagcagaagttgaaagataatcatcagttaacttcaagaatcaggctct       |          |              |                                 | 3027 |
| Db | 2616 | CCCTACATCAAGAGCGGAAGTTGGAAGAACATCATGTGAACCTTTCAAAAACGAGCGCT       |          |              |                                 | 2675 |
| Qy | 3028 | cgctccatttcccttctatctttagacttatttctttagatgagaagaatcagagtgagaagaga |          |              |                                 | 3087 |
| Db | 2676 | CGTCCCTATTCTCTTACTGAGCTTATTTCTTATCCGATGATCAGGAGCAAGGGCA           |          |              |                                 | 2735 |
| Qy | 3088 | gaacctagaataaaacttgtcgaagccttaatgaaaccaaaacttacttctgaaagtgcga     |          |              |                                 | 3147 |
| Db | 2736 | GAACTGCAACACACTTCTGTTCAGGCCAAATGAAGAACCAAGAACTTACTTTTGGAAATGGAG   |          |              |                                 | 2795 |
| Qy | 3148 | catcatatgacccacttaagaatgagtttgcactgcaaaagccttggctatttctctgat      |          |              |                                 | 3207 |

Dh 2796 CATCACATGACACCCAGAGACGATTTTGACTGCAAAAGCCTGGCCCTACTTTTCTGAT 2855  
Qy 3208 gttgacctggaanaagatgtgcactcaaggccctgatttgaaacctctctggctcgcacact 3267  
Dh 2856 GTTGACCTGGAAAAGATGTGCACTGACGCTTGATGGCCCCCTTGATCTGCGCCGCG 2915  
Qy 3268 aacacacgtgaacctgtctcaatgggagcaagtgcagctcagaatctgctctgtttctc 3327  
Dh 2916 AACACCCCTGAACGCTGCTCACGCTGACAACTGACCGTCCAGAAATTTCTCTGTTTTC 2975  
Qy 3328 acctctttgtatgagaccaaagctgtgtactctcaactgaaatatggaagaacctgcag 3387  
Dh 2976 ACTATTTTGTATGAGACAAAAGCTGCTGACTTCTGAAATGTGAAAGAACTCCCG 3035  
Qy 3388 gctccctgcaatatccagatggaagaatcccaactttaagaagaatctgctctcaatga 3447  
Dh 3036 GCCCCTCCACACCTGACAGATGAGAGCCCACTGCAAGAAACATATGCTTCATGGA 3095  
Qy 3448 atcaatggtacataatgagatacactaccttgcttaataatggtcgaagaacaagatt 3507  
Dh 3096 ATCAATGGCTATGTGATGGATCACTCCCTGCTTAGTAACTGAGTCAAGAAAGATC 3155  
Qy 3508 cgatggtatctgctagatgagcagcaatgaaacatccattctatctatctatctagtg 3567  
Dh 3156 CGATGATATCTGCTGATGAGATGGGCAATGAAATATCTCATTTGATTTTACGGA 3215  
Qy 3568 catgtgtcactgtacgaaaaaagaagagataaaatggaactgtacacatctctatca 3627  
Dh 3216 CACGTGTTCACTGTGCGGAAAAAGAGAGATTAATAATGCCGTGACAACTCTATCCG 3275  
Qy 3628 ggtgttttttgagacatgagaaatgtlaccatccaaagctggaatttgcggtgagaaatgc 3687  
Dh 3276 GGTGTCTTTGAGACAGTGAATGCTACCGTCAAAAGTTGGAATTTGGCAATGAAATGC 3335  
Qy 3688 cttaattggcagagctctatcaatgctgtgagatgagacacatttctctgtgtaagcaatag 3747  
Dh 3336 CTGATGTGCGAGACCTGCAAGCTGGATGAGACACACTTCTCTGCTGACAGCAAGAG 3395  
Qy 3748 tgtcagactcccttgagaaatggtctctgacacatlaagattctcagaatcaagcttca 3807  
Dh 3396 TGTGAGGCTCACTGCGAATGGCTTCTGAGCAGCATTTAGATTTTTCAGATCAAGCTTCA 3455  
Qy 3808 ggaacaatggaacagtgagcccaagaagctgacgaactctatctcgatcaatcaat 3867  
Dh 3456 GGACAGATGAGACATGGGCCCCCAAGCTGGCCAGACATTCATTATCCCGATCAATCAAT 3515  
Qy 3868 ggcctggaagcacaagaagagcccttctcttgatcaagtgatcgtttggcacaatgatt 3927  
Dh 3516 GCCTGGAGACCAAGATGCCCATCTCTGGATCAAGTGGATGTTGGACCAATGATGC 3575  
Qy 3928 attcagggcatcaagaaccagggtgcgccgtcagaagttctccagcctctacatctctcag 3987  
Dh 3576 ATTCAAGGCATCATGAGCCAGGTCGCCGTGAGAAAGTTTTCAGGCTCTACATCTCCGAG 3635  
Qy 3888 ttatcatcatgtagtctctgagatgggaagaagtggcagaactctatcaggaatccact 4047  
Dh 3636 TTTATCATCATGTACAGTCTTGACGAGGAGAACTGGCAGAGTTTACCGAAGGAAATTCACAG 3695  
Qy 4048 ggaaccttaatgctcttttttgagaaatgtaattcatctgggataaacaacatctttt 4107  
Dh 3696 GGCACCTTAATGGTCTTTTGGCAATGTGGAGCATCTGGGATTAACACATATATTTT 3755  
Qy 4108 aacctccaatatctgctcgaatacgttttgacccaacacatctatagactctgcagc 4167  
Dh 3756 AACCTCGCATTTGTGGCTCGGTACATCGTTTGCACCCACACATTTACAGCATCCGACAC 3815  
Qy 4168 acctctcagatgagtgagtgaggtctgagcttaaaatgctgcagcatgccaattgggaatg 4227  
Dh 3816 ACTCTTGCGCATGAGATGATGGGCTGATTAATTAACAGTTGCAGCATGCCCCGGGAAATG 3875  
Qy 4228 gagaagtaagaataatcagatgagacagattactgcttcaactctatcttaacatgttt 4287  
Dh 3876 CAGAAATGAAGGATATGAGACTCACAGATCAACGGCTCTCCCTCCACCTAAGCAATATATTT 3935

Qy 4288 gccacctgtctctctcaaaaagctcgacttcaactcctccaagaaggagatgactctgaga 4347  
Dh 3936 GCCACCTGGTCTCTCTCACAAAGCCGACATTCACCTCCAGGGCGGAGCAATGCTCGGGA 3995  
Qy 4348 cctcagtgtaataatccaaaagatggtctgcgaatgagacttccagaagaacaatgaatc 4407  
Dh 3996 CCCCCGTGAGAGCCCAAGAGATGGCTGACAGGTGACCTTCAAAACAGGTGAAGTCT 4055  
Qy 4408 acagagtaactactcaggagtaaatctctgcttaccagatgtatgtgaagagctc 4467  
Dh 4056 ACAGGATCACCAACCCAGGCGCTGAAGTCTCTGCTCAGCAGCATATATGTGAAGAGTTC 4115  
Qy 4468 ctcatctcagaagatcaaatgagatgagctcagtgactctcttcttccaagaatgcaagta 4527  
Dh 4116 CTCGTGTCCAGATGATGAGAGCGCGCCGCTGACCCCTGTTCTTCAGAGAGCGCACAG 4175  
Qy 4528 aaggttttccagggaatcaaatgactctctcaacactgtgtgtgaactctctcagaaccag 4587  
Dh 4176 AAGGTTTTCAAGGCAATCAGGACCTCTCCACCCTGCTGGAAGCGTCTGGACCCCGC 4235  
Qy 4588 ttaactgactcgactcacttgaatcaaccccaagattggtgcaccagattgacctgag 4647  
Dh 4236 CTGTTACGCGCTACTCTGAGATCCACCCACGACTGGGCGCAGCATGCGCTGAGG 4295  
Qy 4648 atggaagttctggtcgtcgaaggcacagacctctactga 4686  
Dh 4296 CTCGAGGTTCTAGATGTGAGGACAGAGATCTCTACTGA 4334

RESULT 13  
US-08-717-294-42  
Sequence 42, Application US/08717294  
Patent No. 6114148  
GENERAL INFORMATION:  
APPLICANT: SEED, BRIAN  
APPLICANT: HAAS, JURGEN  
TITLE OF INVENTION: HIGH LEVEL EXPRESSION OF  
TITLE OF INVENTION: PROTEINS  
NUMBER OF SEQUENCES: 110  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/717,294  
FILING DATE: 20-SEP-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/345001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX:  
INFORMATION FOR SEQ ID NO: 42:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4451 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear

MOLECULE TYPE: cdna  
US-08-717-294-42

Query Match 54.7%; Score 2605.4; DB 3; Length 4451;  
Best Local Similarity 74.8%; Pred. No. 0;  
Matches 3266; Conservative 0; Mismatches 1101; Indels 0; Gaps 0;

QY 328 gccacagaagatactactcgtgctgagtgagtgacatgacagtgagtgacatgacagtgat 387  
DB 84 gccacagccgctctactcgtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 143  
QY 388 ctgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 447  
DB 144 ctgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 203  
QY 448 aacacacagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 507  
DB 204 aacacacagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 263  
QY 508 atgcctgaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 567  
DB 264 atgcctgaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 323  
QY 568 tatgatacagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 627  
DB 324 tatgatacagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 383  
QY 628 gtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 687  
DB 384 gtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 443  
QY 688 agggagagaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 747  
DB 444 agggagagaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 503  
QY 748 ctgagagaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 807  
DB 504 ctgagagaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 563  
QY 808 catgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 867  
DB 564 catgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 623  
QY 868 gaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 927  
DB 624 gaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 683  
QY 928 gtatgagagaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 987  
DB 684 gtatgagagaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 743  
QY 988 gatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1047  
DB 744 gatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 803  
QY 1048 tctctgagagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1107  
DB 804 tctctgagagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 863  
QY 1108 ggcacacacagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1167  
DB 864 ggcacacacagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 923  
QY 1168 catgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1227  
DB 924 catgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 983  
QY 1228 atgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1287  
DB 984 atgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1043  
QY 1288 gaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1347

DB 1044 gaggcttaccgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1103  
QY 1348 gaagaagcggagaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1407  
DB 1104 gaggagcccgaggcttaccgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1163  
QY 1408 gatgataaacctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1467  
DB 1164 gatgataaacctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1223  
QY 1468 acttgggtacatcattcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1527  
DB 1224 acttgggtacatcattcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1283  
QY 1528 gccccgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1587  
DB 1284 gccccgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1343  
QY 1588 aggaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1647  
DB 1344 aggaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1403  
QY 1648 gctatcagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1707  
DB 1404 gctatcagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1463  
QY 1708 ctgttgatattatgaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1767  
DB 1464 ctgttgatattatgaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1523  
QY 1768 actgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1827  
DB 1524 actgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1583  
QY 1828 ttcccaattcccgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1887  
DB 1584 ttcccaattcccgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1643  
QY 1888 ccaactaatcagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1947  
DB 1644 ccaactaatcagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1703  
QY 1948 agagatcagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 2007  
DB 1704 agagatcagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1763  
QY 2008 caaagagagaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 2067  
DB 1764 caaagagagaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1823  
QY 2068 gagaagcgaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 2127  
DB 1824 gagaagcgaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1883  
QY 2128 gtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 2187  
DB 1884 gtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1943  
QY 2188 gttctgataagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 2247  
DB 1944 gttctgataagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 2003  
QY 2248 agcatgtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 2307  
DB 2004 agcatgtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 2063  
QY 2308 aaaaagtgcataagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 2367  
DB 2064 aaaaagtgcataagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 2123  
QY 2368 tcgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 2427

|    |      |   |      |
|----|------|---|------|
| Db | 2124 | TCATATGAGAAACCCCGGCGCTGTGGATTCTTGGGCTGCCACAACAGCGACTTCCGACACCGC     | 2185 |
| Qy | 2428 | ggcattgaccgcgccttactgaaagtttctagttgtgtgacaagaacaccttggltgattatcagag | 2487 |
| Db | 2184 | GGCATGACTGCGCTGCTGTAAGTCTCCAGCTGGCAGCAAAACACCGGGGACTTACTGACAG       | 2243 |
| Qy | 2488 | gaccgttatgaaagatatlttcagcatctgtcgaagtaaaaaaatgcattgcattgaaaccaaga   | 2547 |
| Db | 2244 | GACAGCTAGGAGAGACATCTCCGCTTACTCTCTGTCCAAAGAACACAGCGCATGTGACCGCCGC    | 2303 |
| Qy | 2548 | agctctcccaagattccaagaacacccagcactctggcaaaagcaatttaattgcacccca       | 2607 |
| Db | 2304 | TCCCTTCTCCCAAAACTCCGCCACCCAGCAGCGGTGAGAAGCATTCACGCCACCCCC           | 2363 |
| Qy | 2608 | ccagctcttgaacgcgcatacacaacgggaaataactgcatactctcttcagtcagataaagag    | 2667 |
| Db | 2364 | CCCGTGTCTAAGAGCGCCACCGCCGAGATCACCGGCACACCCCTGCAAGAGCACCAGAG         | 2423 |
| Qy | 2668 | gaaatttgactatgatgataccatacctgaattgaaatgaaaggaagatttggacatttat       | 2727 |
| Db | 2424 | GAGATCTGACTTACACGACACACATTCAGCGTGGAGTGAAGAAAGGAGGACTTCGACATCTAC     | 2483 |
| Qy | 2728 | gattgagatgaaatatcagaacccccgcagctttccaagaagaacagacattattat           | 2787 |
| Db | 2484 | GACGAGAGAGAGAACCAAGAGCCCGCGTCTCTTCCAAAGAAAACCCGCCACTTCTATC          | 2543 |
| Qy | 2788 | gctgcagctggaagagcgtcttggattatggaattgagtatgcctccacatatgttcttaagaac   | 2847 |
| Db | 2544 | GCGCGCGTGGAGCGCTGTGGGACTACGGCATGAGCAGACCCCCACAGTCTCTCGCAC           | 2603 |
| Qy | 2848 | agggtctcaagatgagcaggtgtccctcagttctaagaanaattgtttccaaggaaattacgat    | 2907 |
| Db | 2604 | CGCGCCCAAGCGCGCAGCGCGCCCACTTCAGAGAGGTGGTGTCCAGAGACTTCACCGAC         | 2663 |
| Qy | 2908 | ggctcccttactcaagccctatatcccgctggagaacataatgaacaatttggagctccttggg    | 2967 |
| Db | 2664 | GGCAGCTTACCCAGCGCCCTGTACCGGGGAGGTGAAACAGACACTGTGGCTCTCGCG           | 2723 |
| Qy | 2968 | ccatatataagcagagaaglttgaaagataatcatatgltlaacttcagaatatcagccctc      | 3027 |
| Db | 2724 | CCCTACATCCGGCCCGAGGTGGAGGACMACATCATGTGTGACTTCCGCAACCAAGCCCTCC       | 2783 |
| Qy | 3028 | gctccctatctctctcatctccttagcttatcttctaibagaagaatccagagccaagagca      | 3087 |
| Db | 2784 | CGGCGCTTACTCTTCTACTCTCTCCCTGATACGCTACGAGGAGGACCAAGCGCAGGGCGC        | 2843 |
| Qy | 3088 | gaaccatagaanaaacttgttccaagcttaatgaaacccaacacttactcttggaaagtgcga     | 3147 |
| Db | 2844 | GAGCCCGCGAAGAACTTCGTGAAGCCCAAGACAGACTTAAGACTTCTGTGAAGAGTGCAG        | 2903 |
| Qy | 3148 | catcataatggaccocaataaagatlgatttgaactgcaaaagccttggcttatctctgat       | 3207 |
| Db | 2904 | CACCAATATGGCCCCCACAAGAGAGAGATTCGACGTGCAGAGCGCTTGCGCTTACTTCAGCGAC    | 2963 |
| Qy | 3208 | gtttgaccttgaaaanaaajtgagactcaagcctgaattbgaccctcttgcgtctgccaact      | 3267 |
| Db | 2964 | GTGGAGCTTGAGAGAGACGTGCACAGCGGCTGTACTCGGCCCTGTGTGTGCAACCC            | 3023 |
| Qy | 3268 | aacacacttgaacccctgcctcattggagaacaatgagcatacagaagatttgcctgttttc      | 3327 |
| Db | 3024 | AACACCCCTGAACCCCGCCCAAGGAGGAGCAGTAGTACTGTGCAAGAAATTTGCCCTTCTTC      | 3083 |
| Qy | 3328 | accatcttttgaatgagaaccaaaagcctgttacttcaacttgaanaatatgaaagaanaactgcag | 3387 |
| Db | 3144 | GGCCCCCTGAACCTCCAGATGAGAGATCCACCTTCAGAGAGAACTACCGCTTCACGGCC         | 3203 |
| Qy | 3448 | atccaatgctacataatggaataacactccggcttataatgtgtcgaagtccaagaagt         | 3507 |
| Db | 3084 | ACCATCTTTCGACAGAGACTTAAGACTGTGACTTCCACCGAAGAACTGAGAGCGCAATGTCCGC    | 3143 |
| Qy | 3388 | gtctcccttgaatatccaagatggaagatcccaacttttaagagaaattatgccttccatgca     | 3447 |
| Db | 3144 | GGCCCCCTGAACCTCCAGATGAGAGATCCACCTTCAGAGAGAACTACCGCTTCACGGCC         | 3203 |
| Qy | 3448 | atccaatgctacataatggaataacactccggcttataatgtgtcgaagtccaagaagt         | 3507 |
| Db | 3204 | ATCAACGGGTACTATCTATGTGACACCTCTCCCGCGCTGTGATGTGGCCACGAGACAGCATC      | 3263 |

|    |      |   |      |
|----|------|---|------|
| QY | 3508 | cgatggtatctgctcagcatggtgcagcaatgtaaacatccattctatcttaattcagtgga      | 3567 |
| Db | 3264 | CGCTGGTACCTGCTGTATTATGGGAGCAACGAAGAACTACCAAGCATCTCCACTTTCAGCGCG     | 3323 |
| QY | 3568 | catggttccactctgacgaaaaaagagagatataaaatgacgtgacatgacatctcatca        | 3627 |
| Db | 3324 | CACGTTTTTCACCGTGGCCGAGAAAGAGAGATACAAAGATGGCCCTGTACACACTGTACCCC      | 3383 |
| QY | 3628 | ggtgtttttgagacagtgtgaaatgttcaatccaaagctgtgaaatttggcgggtgaaatgc      | 3687 |
| Db | 3384 | GCGCGTTCGAGACTGTGGAGATGCTCCACGAAGAGCGGGCATCTGGCGCTGGAGATGC          | 3443 |
| QY | 3688 | cttatgtagcgacgtacacatgtcgtggatagagacactttttctgtgtatgaagcaaatg       | 3747 |
| Db | 3444 | CTGATCGCGGAGACCTGTGCAGCGCGGATAGGACCTGTTCTTGTTGATGACGACAAAG          | 3503 |
| QY | 3748 | tgtaagactctcccttgggaatgtgctctctgcagacatttagagatlttcaagatltcaagcttca | 3807 |
| Db | 3504 | TGCCAGACCCCCCTGGGATGTGGCAGCGGCACATCTCGGAGATTCAGATACCGCCAGC          | 3563 |
| QY | 3808 | ggacacataatgacagctgtggccccaagctgtgcagagcttcattatctcgatccaatcat      | 3867 |
| Db | 3564 | GGCCAGTACGGCGAGTGGGCTCCCAACCTTGGCCCGCTCACTTACAGCGCGACGATCTAAC       | 3623 |
| QY | 3868 | gcttgtagcacaagagagccctttcttggataagtgatctgttgcagcaatgata             | 3927 |
| Db | 3624 | GCTGTGTGACCAAGAGAGCCCTTCTCTGATGATCAAGTGGACCTGTGGCCCCCATGATC         | 3683 |
| QY | 3928 | atleaagcatalcaagaccsaaggtgcocgtcagaagtcttccagctccatactctcag         | 3987 |
| Db | 3684 | ATCCACGGCATTCAGACCCAGCGGGCGCCGACAAAGTTTACGACGCTGTATCGACCGAC         | 3743 |
| QY | 3988 | tttatcatcatgataagttcttggatggaaaagatggcagacttaccgagaaattccact        | 4047 |
| Db | 3744 | TTTCATCATCATGTACTCTCTTAGACGCGAAGATGGCAACCTACC GGCGCAACGACACC        | 3803 |
| QY | 4048 | ggaaccttaatgtgctctcttcttggcaatgttgattatcctggatataaacacaataatltt     | 4107 |
| Db | 3804 | GGCACCTCGATGGTGTCTTTCGGCAAGTGGACAGCAGCGGATACAGCACACATCTTC           | 3863 |
| QY | 4108 | aaacctccaattatgtctgatatacatcgttttcaccacaacatcatatagaatctgcagc       | 4167 |
| Db | 3864 | AACCCCCCATCATGCGCCGCTATCATCGGCTCACCCCAACCACTACAGCATCCGACGC          | 3923 |
| QY | 4168 | actcttcgcatgagatgtgataggcgctgtgatttaaatagttgcagatgcatttggaaatg      | 4227 |
| Db | 3924 | ACCGTGGCCATGTGAGTGATGGGCTCGACCTGAACAGTGGCAGATGGCCCTGGGCATG          | 3983 |
| QY | 4228 | gagagttaagcaatatcaagatgtcagagatctgcttcatctactacttaaccaatagtltt      | 4287 |
| Db | 3984 | GAGACCAAGGGCATACAGCGACGCCCAATACCGCTCCAGTACTTCCACCAACTGTTC           | 4043 |
| QY | 4288 | gcaacctgtctcttccaagaagctgcagatctcactccaaggaggaatgtaatgcctggaga      | 4347 |
| Db | 4044 | GCCACTCTGGAACCCCGACAGAGCCCGGCTCTACCTGAGAGGCCGACGAAACCCCTGGGCG       | 4103 |
| QY | 4348 | cctcaagtgtaaatltcaaaaagatgtgctgcagatgagacttccagaagaacatlgaaatgc     | 4407 |
| Db | 4104 | CCCCGGTGAACACCCCAAGAGATGGCTCTCAGGTGGATCTTCCAAACCAACCTGAAAGGTG       | 4163 |
| QY | 4408 | acagagagtaactactctcagggagtaaaatctctgtctaccagaatgtatgtgaagagctc      | 4467 |
| Db | 4164 | ACTGGCGTACACCCAGGGCGTCAAGAGCTCTTACCAAGCATGTGACTGAAAGAGTTC           | 4223 |
| QY | 4468 | ctcaatctccaacgataaagatgtgcatactggaactctcttttttccaagatgtgcaaaatga    | 4527 |
| Db | 4224 | CTGATTCACACACACAGAGAGCGGCAACATGTGACCTGTCTTCCAAAGCGCAAGGTG           | 4283 |
| QY | 4528 | aaggttttccagggaataatcaagactcctctcaacactgtgtgtgaactctctagaccacag     | 4587 |
| Db | 4284 | AAGGTGTTCAGGGCAACAGAGAGTTCATCACCGCTGTGTAAACAGCTGTGACCCCCCCC         | 4343 |







|    |      |  |      |
|----|------|--|------|
| QY | 270  | catgcaaaatagagctctccaccctgctctctctctctgaccttctgacatctctgcttaagtc   | 329  |
| Db | 150  | catgcaaaatagagctctccaccctgctctctctctctgaccttctgacatctctgcttaagtc   | 209  |
| QY | 330  | caccgaagaataactactggtggtgagcgtgatacgttgaactatgaggaataatgcaatgacct  | 389  |
| Db | 210  | caccgaagaataactactggtggtgagcgtgatacgttgaactatgaggaataatgcaatgacct  | 269  |
| QY | 390  | cagtgagctgaccttgagacgaagaattctctctctctctctctctctctctctctctctctctct | 449  |
| Db | 270  | cagtgagctgaccttgagacgaagaattctctctctctctctctctctctctctctctctctctct | 329  |
| QY | 450  | caacctcagtcgtgtaacaaaagaacctcgttttgttagaattcaacgtgttcacacttcaacat  | 509  |
| Db | 330  | caacctcagtcgtgtaacaaaagaacctcgttttgttagaattcaacgtgttcacacttcaacat  | 389  |
| QY | 510  | cgtctaacgcaaaagccacccttgatggctctgctagtgctctacacatccagcgtgaagtta    | 569  |
| Db | 390  | cgtctaacgcaaaagccacccttgatggctctgctagtgctctacacatccagcgtgaagtta    | 449  |
| QY | 570  | tgaatacaatgtctctcaacttaagaacatggctctccactctgtaagctctatgctgct       | 629  |
| Db | 450  | tgaatacaatgtctctcaacttaagaacatggctctccactctgtaagctctatgctgctgct    | 509  |
| QY | 630  | tgtgtatctaccctacggaagaacctctctctgaggaagcttgatatgtatgatacaacagccaag | 689  |
| Db | 510  | tgtgtatctaccctacggaagaacctctctctgaggaagcttgatatgtatgatacaacagccaag | 569  |
| QY | 690  | ggagaagaagaatgataaagctctccctggttgaaacataatgctctgagagctct           | 749  |
| Db | 570  | ggagaagaagaatgataaagctctccctggttgaaacataatgctctgagagctctgagagctct  | 629  |
| QY | 750  | gaaagaagaatgtccaatgtgacctctgaacacagctgacctctactactaatatctcttca     | 809  |
| Db | 630  | gaaagaagaatgtccaatgtgacctctgaacacagctgacctctactactaatatctcttctca   | 689  |
| QY | 810  | tgtgaacctgtgtaaaagaacttgaaattcaagcctcatcttgagccctactagttgttagaga   | 869  |
| Db | 690  | tgtgaacctgtgtaaaagaacttgaaattcaagcctcatcttgagccctactagttgtgttagaga | 749  |
| QY | 870  | agggaagctgtgccaagaaaagacacagacctgtgacaaattatatactactcttctgct       | 929  |
| Db | 750  | agggaagctgtgccaagaaaagacacagacctgtgacaaattatatactactcttctgctgct    | 809  |
| QY | 930  | atttgatgaagaagaagttgtgacctctagaacaaaagaactccttgatgacgaataagga      | 989  |
| Db | 810  | atttgatgaagaagaagttgtgacctctagaacaaaagaactccttgatgacgaataagga      | 869  |
| QY | 990  | tgtctcatctctctgagccctgagccctaaagaagcacagctcaatgtgtatgtataacagctc   | 1049 |
| Db | 870  | tgtctcatctctctgagccctgagccctaaagaagcacagctcaatgtgtatgtataacagctc   | 929  |
| QY | 1050 | tcctgcaggctctgtatgagatgccaagaaatcagctcatcttgagcatgagatggaatggg     | 1109 |
| Db | 930  | tcctgcaggctctgtatgagatgccaagaaatcagctcatcttgagcatgagatggaatggg     | 989  |
| QY | 1110 | caccactctctgaagtgtcatctcaatatctccctgaaggtacacacatctctctgtgaagaacca | 1169 |
| Db | 990  | caccactctctgaagtgtcatctcaatatctccctgaaggtacacacatctctctgtgaagaacca | 1049 |
| QY | 1170 | tcgcaagcgcctcttggaatactcgcaacataaactctctactgagctccaacacctctgat     | 1229 |
| Db | 1050 | tcgcaagcgcctcttggaatactcgcaacataaactctctactgagctccaacacctctctgat   | 1109 |
| QY | 1230 | ggaccttggaaagtttctactggtttgtgcatatctcttccacacacatgataatgagcatgga   | 1289 |
| Db | 1110 | ggaccttggaaagtttctactggtttgtgcatatctcttccacacacatgataatgagcatgga   | 1169 |
| QY | 1290 | agcttatgtcaaaagtgaacagcgtgtccagaaggaaaccccaactacgaatgaaataatgta    | 1349 |
| Db | 1170 | agcttatgtcaaaagtgaacagcgtgtccagaaggaaaccccaactacgaatgaaataatgta    | 1229 |

|    |      |  |      |
|----|------|--|------|
| QY | 1350 | agaagcggaaagactatgaatgatctctactgatctctgaagatggtatgcttgta         | 1409 |
| Db | 1230 | AGAAACGGGAACACATATGATGATGATCTTACTGATCTTGAAATGATGTGGTCAGGTTTGA    | 1289 |
| QY | 1410 | tgaatgaacactctccctcccttaaccaaatctgactgaatttgcagaagaacatccctaaac  | 1469 |
| Db | 1290 | TGATGACACACTCTCCCTCTTTATCCAAATTCGCTCAGTTGGCCAAAGAACATCTTAAAC     | 1349 |
| QY | 1470 | ttgggtataactatactgcctgtgaagggagagacttggactatgctccctatgctctgc     | 1529 |
| Db | 1350 | TTGGGTTACATTTACATTTGCTGCTGTAAGAGAGAGACCTGGACATATGCTCCTTATGTCCTGC | 1409 |
| QY | 1530 | ccccgaatgaacgaagtataaaagatcaatttgaacaaatggccctcagcggatgtgtag     | 1589 |
| Db | 1410 | CCCCGATGACAGAAAGTTATATAAATCAATTTTGAACAATGGCCCTCACGCGATGTTAG      | 1469 |
| QY | 1590 | gaagatcaaaaaaagtcggattttatggtatcacaaatgaaacctttaagactcgtgaac     | 1649 |
| Db | 1470 | GAAGATCAAAAAAATCCGATTTATGGCATTCACAGATGAACCTTTAAACATCGTGAAGC      | 1529 |
| QY | 1650 | tattcaagaatgaatcagaatcttggagacttacttiaatgggaaagtggagaacact       | 1709 |
| Db | 1530 | TATTTCAGATTAATTCAGGAATCTTTGGACCTTTACTTTATGGGAAGTTGGAGACACACT     | 1589 |
| QY | 1710 | gttgatataatlttaagaatcaagcaagcagaccataataacatctaccctcacaagaaatcac | 1769 |
| Db | 1590 | GTTGATTTATTTTAAAGATCAAGCAACAGACATTAATCAATCTACCTCACGGAATCAC       | 1649 |
| QY | 1770 | tgaatgcgctcccttctgatactcaagggagattaccaaaagtgtaaaaaatttgaagatct   | 1829 |
| Db | 1650 | TGATTCGCTCCTTTGATTCMAAGAAATTTCCAAAAGGTGTAAACATTTTGAAGATTT        | 1709 |
| QY | 1830 | tccaaattctgcagagaagaatattccaataataatgaaagatgactgtgaagaatctgggac  | 1889 |
| Db | 1710 | TCCAAATTCTGCCAGAGAAATATTCCAATTTAAATGGAAGATGACTGTGGAAGATGGGCC     | 1769 |
| QY | 1890 | aactaaatcagaatccctcgtgtgcttgacccgctatctactagtcttctgtaatatggaag   | 1949 |
| Db | 1770 | AACTAAATCAAGATCTCTGGGGCCMGACCCGCTATTTACTGATTTCGTTAATATGGAAG      | 1829 |
| QY | 1950 | agatctagcttcagactcaatttggccctctccctcaatctgctatacaagaatctgtgatca  | 2009 |
| Db | 1830 | AGATCTACCTTTACAGAGCTCATTTGGCCCTCCCACTCTCTCAAAAGATCTGTATGATCA     | 1889 |
| QY | 2010 | aagagaaacagaataatgtcacagaagaagaaatgtcatccgtttctcgtatatttgatga    | 2069 |
| Db | 1890 | AAGAGAAACACAGTAATGTTCAGACAAAGAGGAATGTCACTCTGTTTCTGTATTTGAAGA     | 1949 |
| QY | 2070 | gaaccgaagctgtgacctcacaagaaatalacaagctcttcccacaatccacgttgaagt     | 2129 |
| Db | 1950 | GAACCGAAAGCTGTACTCTCACAGAGAAATATACAGCGTTTCTCCCATTCACAGCTGGAGT    | 2009 |
| QY | 2130 | gcaagcttgaagatccagagttccaagctccaacatcatatgacagatcaatcatgtctatgt  | 2189 |
| Db | 2010 | GCAGCTTGAGATTCACAAGTTTCCAAAGCTTCCCAATCATCATGACACATCAATATGGCTATGT | 2069 |
| QY | 2190 | ttttgataagtttgaagtgtcgaattgttgtttgaaatggaggggtgaatactgtatcatctaa | 2249 |
| Db | 2070 | TTTTGATAGTTTGCAGTTTGCAGTTGTTTGGCATGAGGGGTACTGTGATCAATTCTAAG      | 2129 |
| QY | 2250 | catlggaacacagaacttactctctctgtctctctctctctgtatataactcttaaacacaa   | 2309 |
| Db | 2130 | CATTGGAGCACAGACTACCTTCCCTTCTGTCTTCTCTGTGATATTACCTTCAAAACAA       | 2189 |
| QY | 2310 | aaatgctctatgaagaacactcacccctatcccatctcagaagaactgtctctcaatgtc     | 2369 |
| Db | 2190 | AATGCTGTATGAAGACACACTTCACCTATTTCCATTTCTCAGAGAAACTGTCTTCATGTCTC   | 2249 |
| QY | 2370 | gatggaaaaacccagttctatgatttttgggggtggcaacaactagaactttggaacaagg    | 2429 |
| Db | 2250 | GATGGAANAACCCAGGTTATGGAATTTCTGGGGTGTCCACAACCTACGACTTTTCCGAACAGAG | 2309 |
| QY | 2430 | catgacgcctctactgaaggtttctcagttgtgtaacaagaacactgtgtatattacagaga   | 2489 |

Db 2310 |||||CATGACCGCCTTACTGAGGTTTCTAGTTGTGACAAGAACACTGGTGATTATTACGAGGA 2369  
Oy 2490 cagttatgaagataatlcagcatacttgctgaataaacaatlgccattgaaccaagaag 2549  
Db 2370 CAGTTATGAAGATATTTCAGCATACTTCTGAGTAAACAAATGCCATTGAACCAAGANG 2429  
Oy 2550 ctctcccaagaatccaagacacccctagcactagcnaaagaatltaaatgccccca 2607  
Db 2430 CTCTCCCAAGATTCAAGACACCCCTAGCAGTAAAGCAATTTAATGCCACCACA 2487

Search completed: November 18, 2001, 05:07:44  
job time: 36066 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2001, 05:14:47 ; Search time 919.72 seconds  
(without alignments)  
3253.111 Million cell updates/sec

Title: US-09-689-430-1-COPY\_150\_4914  
Perfect score: 4765  
Sequence: 1 cctcttctaagtaacagta.....ttgggtcgtttgttcgcgac 4765

Scoring table:  
IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%  
Listing first 45 summaries

Database :  
1: N.Geneseq\_0601:\*  
2: /SIDSR/gcgdata/geneseq/geneseq/NA1980.DAT:\*  
3: /SIDSR/gcgdata/geneseq/geneseq/NA1981.DAT:\*  
4: /SIDSR/gcgdata/geneseq/geneseq/NA1982.DAT:\*  
5: /SIDSR/gcgdata/geneseq/geneseq/NA1983.DAT:\*  
6: /SIDSR/gcgdata/geneseq/geneseq/NA1984.DAT:\*  
7: /SIDSR/gcgdata/geneseq/geneseq/NA1985.DAT:\*  
8: /SIDSR/gcgdata/geneseq/geneseq/NA1986.DAT:\*  
9: /SIDSR/gcgdata/geneseq/geneseq/NA1987.DAT:\*  
10: /SIDSR/gcgdata/geneseq/geneseq/NA1988.DAT:\*  
11: /SIDSR/gcgdata/geneseq/geneseq/NA1989.DAT:\*  
12: /SIDSR/gcgdata/geneseq/geneseq/NA1990.DAT:\*  
13: /SIDSR/gcgdata/geneseq/geneseq/NA1991.DAT:\*  
14: /SIDSR/gcgdata/geneseq/geneseq/NA1992.DAT:\*  
15: /SIDSR/gcgdata/geneseq/geneseq/NA1993.DAT:\*  
16: /SIDSR/gcgdata/geneseq/geneseq/NA1994.DAT:\*  
17: /SIDSR/gcgdata/geneseq/geneseq/NA1995.DAT:\*  
18: /SIDSR/gcgdata/geneseq/geneseq/NA1996.DAT:\*  
19: /SIDSR/gcgdata/geneseq/geneseq/NA1997.DAT:\*  
20: /SIDSR/gcgdata/geneseq/geneseq/NA1998.DAT:\*  
21: /SIDSR/gcgdata/geneseq/geneseq/NA1999.DAT:\*  
22: /SIDSR/gcgdata/geneseq/geneseq/NA2000.DAT:\*  
23: /SIDSR/gcgdata/geneseq/geneseq/NA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1          | 4413   | 92.6        | 4629   | 20    | AAH88293    |
| 2          | 4411.6 | 92.6        | 9354   | 18    | AAH73164    |
| 3          | 4355.4 | 91.3        | 4629   | 16    | AAH76016    |
| 4          | 4351.2 | 91.3        | 4670   | 19    | AAV23339    |
| 5          | 4330.8 | 90.9        | 4999   | 21    | AAH00122    |
| 6          | 4325.6 | 90.8        | 9164   | 20    | AAH82259    |
| 7          | 4321.6 | 90.7        | 4832   | 19    | AAV19581    |
| 8          | 4321.6 | 90.7        | 4832   | 19    | AAV15338    |
| 9          | 4319   | 90.6        | 12445  | 21    | AAH49232    |
| 10         | 4259.4 | 89.4        | 11933  | 21    | AAH00121    |
| 11         | 4173   | 87.6        | 4275   | 10    | AAH90654    |

|    |        |      |       |    |          |                     |
|----|--------|------|-------|----|----------|---------------------|
| 12 | 4171.6 | 87.5 | 4275  | 9  | AAH80446 | Modified factor VI  |
| 13 | 4170   | 87.5 | 4272  | 9  | AAH80447 | Modified factor VI  |
| 14 | 4161   | 87.3 | 4373  | 20 | AAH82258 | Beta-domain delete  |
| 15 | 4152.2 | 87.1 | 4830  | 9  | AAH81544 | Human Factor VIII-  |
| 16 | 4108.8 | 86.2 | 4545  | 9  | AAH80444 | Modified factor VI  |
| 17 | 3748   | 78.7 | 5035  | 18 | AAH69811 | Factor VIII-dB695-  |
| 18 | 3708.6 | 77.8 | 4616  | 9  | AAH81545 | Human Factor VIII-  |
| 19 | 3678   | 77.2 | 5094  | 21 | AAA49231 | DNA construct HSO/  |
| 20 | 3189.8 | 66.9 | 4334  | 20 | AAH1439  | Homo sapiens facto  |
| 21 | 3189.8 | 66.9 | 4334  | 20 | AAH91195 | Porcine factor VII  |
| 22 | 3189.8 | 66.9 | 4334  | 22 | AAH90541 | CDNA encoding porc  |
| 23 | 2605.4 | 54.7 | 4451  | 19 | AAH23288 | Synthetic human Fa  |
| 24 | 2336.8 | 49.1 | 11846 | 20 | AAH82261 | Factor VIII protei  |
| 25 | 2336.4 | 49.0 | 9009  | 14 | AAH05185 | Human Factor VIII   |
| 26 | 2336.4 | 49.0 | 9009  | 18 | AAH61548 | Factor VIII:C (Arg  |
| 27 | 2336.4 | 49.0 | 9009  | 19 | AAH25810 | Human factor VIII   |
| 28 | 2336.4 | 49.0 | 9009  | 20 | AAH91162 | Human factor VIII   |
| 29 | 2336.4 | 49.0 | 9009  | 22 | AAH90508 | Human factor VIII   |
| 30 | 2334.8 | 49.0 | 6300  | 17 | AAH03571 | Factor-VIII CDNA.   |
| 31 | 2334.8 | 49.0 | 8241  | 9  | AAH81439 | Factor VIII CDNA. 1 |
| 32 | 2334.8 | 49.0 | 8241  | 9  | AAH81096 | CDNA sequence enco  |
| 33 | 2334.8 | 49.0 | 8975  | 6  | AAH50054 | Human factor VIII   |
| 34 | 2334.8 | 49.0 | 8975  | 21 | AAH38604 | Human full-length   |
| 35 | 2334.8 | 49.0 | 9029  | 22 | AAH60309 | Human factor VIII   |
| 36 | 2334.8 | 49.0 | 9068  | 19 | AAH15359 | Human factor VIII   |
| 37 | 2334.8 | 49.0 | 9080  | 19 | AAH19580 | Human factor III e  |
| 38 | 2333.8 | 49.0 | 7053  | 18 | AAH51357 | Factor VIII:C codi  |
| 39 | 2333.2 | 49.0 | 8967  | 17 | AAH31031 | Factor-VIII full-I  |
| 40 | 2333.2 | 49.0 | 8967  | 22 | AAH87526 | Human factor VIII   |
| 41 | 2333.2 | 49.0 | 9009  | 19 | AAH18884 | Homo sapiens facto  |
| 42 | 2332.6 | 48.9 | 7056  | 15 | AAH06615 | Sequence of human   |
| 43 | 2323.6 | 48.8 | 7440  | 7  | AAH60689 | Sequence encoding   |
| 44 | 2313.8 | 48.6 | 12022 | 20 | AAH82260 | Factor VIII protei  |
| 45 | 2272.4 | 47.7 | 7440  | 6  | AAH50375 | DNA sequence enco   |

## ALIGNMENTS

|          |  |
|----------|--|
| RESULT 1 |  |
| AAH88293 | standard; DNA: 4629 BP.  |
| XX       |  |
| AC       | AAH88293;  |
| XX       |  |
| DT       | 24-SEP-1999 (first entry)  |
| XX       |  |
| DE       | Human Factor VIII with B domain deleted cDNA.                          |
| XX       |  |
| KW       | Adenoviral vector; Factor VIII; Factor IX; clotting factor; treatment; |
| KW       | haemostatic; haemophilia A; haemophilia B; gene therapy; ss.           |
| XX       |  |
| OS       | Homo sapiens.  |
| XX       |  |
| PN       | US935935-A.  |
| XX       |  |
| PD       | 10-AUG-1999.   |
| XX       |  |
| PF       | 07-JUN-1995; 95US-0484891.   |
| XX       |  |
| PR       | 07-JUN-1995; 95US-0484891.   |
| PR       | 10-JUN-1993; 93US-0074920.   |
| PR       | 25-MAR-1994; 94US-0218335.   |
| XX       |  |
| PA       | (GENE-) GENETIC THERAPY INC.   |
| XX       |  |
| PI       | Connelly S, Kaleko M, Smith T;   |
| XX       |  |
| DR       | WPI; 1999-457617/38.   |
| PT       | Adenoviral vectors useful for treating hemophilia                      |
| XX       |  |
| PS       | Example 1; Column 53-58; 90pp; English.                                |

XX This invention describes novel adenoviral vectors comprising at least  
CC one DNA sequence encoding a clotting factor (Factor IX or Factor VIII).  
CC The vectors of the invention have haemostatic activity. The vectors are  
CC useful for the treatment of hemophilia A or hemophilia B by gene therapy.  
CC This sequence represents human Factor VIII cDNA which has the B domain  
CC deleted.  
XX

SQ Sequence 4629 BP; 1319 A; 1050 C; 1015 G; 1245 T; 0 other;

Query Match 92.6% Score 4413; DB 20; Length 4629;  
Best Local Similarity 99.8% Pred. No. 0;  
Matches 4419; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 271 atgcaaatagagctctcactgtctctctctgtgctcttgcagttcgtcttgtagtgc 330  
DB 1 atgcaaatagagctctcactgtctctctctgtgctcttgcagttcgtcttgtagtgc 60  
QY 331 accagaatactactcctgtgtgctgagtgaaactgtcatgtgagctatgtcaaatgtatctc 390  
DB 61 acaagaagatactactcctgtgtgctgagtgaaactgtcatgtgagctatgtcaaatgtatctc 120  
QY 391 ggtgagctgctgtgtgagcgcgaagaattcctcctcgtagtgccaaaattcttccatcaac 450  
DB 121 ggtgagctgctgtgtgagcgcgaagaattcctcctcgtagtgccaaaattcttccatcaac 180  
QY 451 acctcagctgtgtacaataaagactgtgttgtaaatcagtggtcaccttccaacatc 510  
DB 181 acctcagctgtgtacaataaagactgtgttgtaaatcagtggtcaccttccaacatc 240  
QY 511 gctaaagcgaagccacctgtgagtgtctgctagtcctaccatccagctgaagtttat 570  
DB 241 gctaaagcgaagccacctgtgagtgtctgctagtcctaccatccagctgaagtttat 300  
QY 571 gatacagtgctaatctactaactaaagacatggtctccactcgtcagctcctaactgtt 630  
DB 301 gatacagtgctaatctactaactaaagacatggtctccactcgtcagctcctaactgtt 360  
QY 631 ggtgtaccctactctgaaagctctcgaaggagctgaatactgaatcgaaccagtcataaagg 690  
DB 361 ggtgtaccctactctgaaagctctcgaaggagctgaatactgaatcgaaccagtcataaagg 420  
QY 691 gagaagaagaatgataaagctctccctgtgtggaagccaatactgctgcgaagctctg 750  
DB 421 gagaagaagaatgataaagctctccctgtgtggaagccaatactgctgcgaagctctg 480  
QY 751 aagagaagaatgctcaatggtcctcgaaccactgtgcttactactaatcttcttctat 810  
DB 481 aagagaagaatgctcaatggtcctcgaaccactgtgcttactactaatcttcttctat 540  
QY 811 gtgagacctgttaaaagactgaaatctcagcctcatctgagccctactactatgtatgaa 870  
DB 541 gtgagacctgtgtaaagactgaaatctcagcctcatctgagccctactactatgtatgaa 600  
QY 871 gggagctctgcccagaagaagaacagacactctgcacaaattatactactttgtgtga 930  
DB 601 gggagctctgcccagaagaagaacagacactctgcacaaattatactactttgtgtga 660  
QY 931 ttgtgagaagaagaagaattgtgactcagaagaagaagaactcctgtatgcagaataggat 990  
DB 661 ttgtgagaagaagaagaattgtgactcagaagaagaagaactcctgtatgcagaataggat 720  
QY 991 gctgcatctctcgggctctggtcctaaatgacacagtcacatgtgtatgtataacaggtct 1050  
DB 721 gctgcatctctcgggctctggtcctaaatgacacagtcacatgtgtatgtataacaggtct 780  
QY 1051 ctgcccaggtctgattggaatgcacacagaagaatcagtcattatgcatgtgattggaatgggc 1110  
DB 781 ctgcccaggtctgattggaatgcacacagaagaatcagtcattatgcatgtgattggaatgggc 840  
QY 1111 accaccctgaatgtgcctcaataattcctcgaagaagtcacacattcttgtaggaacat 1170  
DB 1111 accaccctgaatgtgcctcaataattcctcgaagaagtcacacattcttgtaggaacat 1170

DB 841 accactcctgaatgtgcctcaataattcctcgaagaagtcacacattcttgtaggaacat 900  
QY 1171 cgcacaggtctccttggaatctgcaccaataacttcttactgtctcaaacactttagt 1230  
DB 901 cgcacaggtctccttggaatctgcaccaataacttcttactgtctcaaacactttagt 960  
QY 1231 gaacttgacagtttctactgtttttgtcatatctcttcccaacaatgcatgtgaa 1290  
DB 961 gaacttgacagtttctactgtttttgtcatatctcttcccaacaatgcatgtgaa 1020  
QY 1291 gcttatgtcaagaatgagacgtctgtccagaaggaaacccacagaaatgaaaataatgaa 1350  
DB 1021 gcttatgtcaagaatgagacgtctgtccagaaggaaacccacagaaatgaaaataatgaa 1080  
QY 1351 gaagcggagaagctatgatatgatacttactgtattcgtgaatgagtggtgcaggttgat 1410  
DB 1081 gaagcggagaagctatgatatgatacttactgtattcgtgaatgagtggtgcaggttgat 1140  
QY 1411 gatgacaactctcctcttctatccaaattcgtcagttgccaagaagcatcctaaact 1470  
DB 1141 gatgacaactctcctcttctatccaaattcgtcagttgccaagaagcatcctaaact 1200  
QY 1471 tgggtacattacattgctgctggaaggagagagctgggagactatgtctcctatgtctgcc 1530  
DB 1201 tgggtacattacattgctgctggaaggagagagctgggagactatgtctcctatgtctgcc 1260  
QY 1531 ccgcatgacagaagttataaagaatcatattgacaatgtgcccctcagcgatgtgag 1590  
DB 1261 ccgcatgacagaagttataaagaatcatattgacaatgtgcccctcagcgatgtgag 1320  
QY 1591 aagtaacaaaagaatccgaattatgtgcatacacagatgaacactttagactgtgaagct 1650  
DB 1321 aagtaacaaaagaatccgaattatgtgcatacacagatgaacactttagactgtgaagct 1380  
QY 1651 attcagcagatcaggaactcttggaacttacttattatgaggaggttgagacacatg 1710  
DB 1381 attcagcagatcaggaactcttggaacttacttattatgaggaggttgagacacatg 1440  
QY 1711 ttgattatatttaagaatcaacgaagcagacacataacatctaccctcaggaatcat 1770  
DB 1441 ttgattatatttaagaatcaacgaagcagacacataacatctaccctcaggaatcat 1500  
QY 1771 gatgtcgtctcttgatctcaaggagatctaccaaaagtgtaaacatttgaagatttt 1830  
DB 1501 gatgtcgtctcttgatctcaaggagatctaccaaaagtgtaaacatttgaagatttt 1560  
QY 1831 ccaattctgcaggaagaataattcaaatataatgagcagtgagatgagatgtggcca 1890  
DB 1561 ccaattctgcaggaagaataattcaaatataatgagcagtgagatgagatgtggcca 1620  
QY 1891 actaatcagatctcctggtgctgacccgctatactactagtttgcgttaatalatgagaga 1950  
DB 1621 actaatcagatctcctggtgctgacccgctatactactagtttgcgttaatalatgagaga 1680  
QY 1951 gatctagcttcaagacatcatgtgcccctctcctaactctgtcacaagaagaatctgtatgaa 2010  
DB 1681 gatctagcttcaagacatcatgtgcccctctcctaactctgtcacaagaagaatctgtatgaa 1740  
QY 2011 agaggaacacagatcaatgtcagacaagaagatgtcaatcctgttttctgtatgtgag 2070  
DB 1741 agaggaacacagatcaatgtcagacaagaagatgtcaatcctgttttctgtatgtgag 1800  
QY 2071 aacggaagctgttacctacacagagaatalataaacagcttctccccaatccagctgtgagtg 2130  
DB 1801 aacggaagctgttacctacacagagaatalataaacagcttctccccaatccagctgtgagtg 1860  
QY 2131 cagcttgagatccagagttcccaagctccaacatcatatgacagcatcaatgtgtatgtt 2190  
DB 1861 cagcttgagatccagagttcccaagctccaacatcatatgacagcatcaatgtgtatgtt 1920  
QY 2191 ttgtatagttgcaggtgttcaagttgtgttgatgaatgagtgacatcgtgtatcatcnaag 2250  
DB 1921 ttgtatagttgcaggtgttcaagttgtgttgatgaatgagtgacatcgtgtatcatcnaag 1980

|    |      |   |      |
|----|------|---|------|
| QY | 2251 | attggaagcaagcgtactccttcccttctgtctcttctcttgatataccctcaacaacaa        | 2310 |
| Db | 1981 | attggagcaacagactgaactctcttctgtctcttctcttgatataccctcaacaacaa         | 2040 |
| QY | 2311 | atggtctatgaaagacacactcaacctatcccatctctcaggagaaactgltctcatgctg       | 2370 |
| Db | 2041 | atggtctatgaaagacacactcaacctatcccatctctcaggagaaactgltctcatgctg       | 2100 |
| QY | 2371 | atggaaaaccccggtctatagatcttcggggtgcacaaactcgaacttcgaacttcggaaacagagc | 2430 |
| Db | 2101 | atggaaaaccccggtctatagatcttcggggtgcacaaactcgaacttcgaacttcggaaacagagc | 2160 |
| QY | 2431 | atgacgcgcttactgtaggtttcttaagtgtgacaagaacactggtatattatcagggac        | 2490 |
| Db | 2161 | atgacgcgcttactgtaggtttcttaagtgtgacaagaacactggtatattatcagggac        | 2220 |
| QY | 2491 | agttatgaaagatatctcagcatactgtctgtagtaaaaaaatgcatctgaaaccaagaagc      | 2550 |
| Db | 2221 | agttatgaaagatatctcagcatactgtctgtagtaaaaaaatgcatctgaaaccaagaagc      | 2280 |
| QY | 2551 | ttctcccaaatctcaagacacccctagcactagcagaagcaatttatctgaccacca           | 2610 |
| Db | 2281 | ttctcccaaatctcaagacacccctagcactagcagaagcaatttatctgaccacca           | 2340 |
| QY | 2611 | gtcttgaagcgcatacgaacggggaataaactcgtactactcttcagtcagatcgaagaa        | 2670 |
| Db | 2341 | gtcttgaagcgcatacgaacggggaataaactcgtactactcttcagtcagatcgaagaa        | 2400 |
| QY | 2671 | attgactatgatatcatatcagtttgaaatgaaaggaagatttgacatttatgat             | 2730 |
| Db | 2401 | attgactatgatatcatatcagtttgaaatgaaaggaagatttgacatttatgat             | 2460 |
| QY | 2731 | ggggatgaaaatccagagcccccgcgaacttccaagaagaacaacgacatttatgct           | 2790 |
| Db | 2461 | ggggatgaaaatccagagcccccgcgaacttccaagaagaacaacgacatttatgct           | 2520 |
| QY | 2791 | gcaagtggaaaggtcctctgggatatctgggacgtagtagctccccaatgcttcgaagaacag     | 2850 |
| Db | 2521 | gcaagtggaaaggtcctctgggatatctgggacgtagtagctccccaatgcttcgaagaacag     | 2580 |
| QY | 2851 | gtctcagagtgagagtgctccctcagttccaagaagttgttttccaaggaatttactgatgac     | 2910 |
| Db | 2581 | gtctcagagtgagagtgctccctcagttccaagaagttgttttccaaggaatttactgatgac     | 2640 |
| QY | 2911 | tccttactcagcccttataccgttgagagactaaatggaacatttgggaccccggggca         | 2970 |
| Db | 2641 | tccttactcagcccttataccgttgagagactaaatggaacatttgggaccccggggca         | 2700 |
| QY | 2971 | tatataagagcagaagtgtgaagataatcatatgtagtaacttcgaagaatcaggccttcgt      | 3030 |
| Db | 2701 | tatataagagcagaagtgtgaagataatcatatgtagtaacttcgaagaatcaggccttcgt      | 2760 |
| QY | 3031 | cccttctcttatattctgagccttatcttcttagaggaagatcaggggaaggaagcaaga        | 3090 |
| Db | 2761 | cccttctctctcatctcagccttatcttcttagaggaagatcaggggaaggaagcaaga         | 2820 |
| QY | 3091 | ccttagaaaaaacttgttccaagccttaatgaaaccaaacttacttggaaatgacaacat        | 3150 |
| Db | 2821 | ccttagaaaaaacttgttccaagccttaatgaaaccaaacttacttggaaatgacaacat        | 2880 |
| QY | 3151 | catatggcaccacataaagaatgagatttgactgcgaagccttggcttatctctctgatt        | 3210 |
| Db | 2881 | catatggcaccacataaagaatgagatttgactgcgaagccttggcttatctctctgatt        | 2940 |
| QY | 3211 | gacccgtgaaaaagatgtgacccaagcctgcatgtgaccccttcgtctcggccaactaac        | 3270 |
| Db | 2941 | gacccgtgaaaaagatgtgacccaagcctgcatgtgaccccttcgtctcggccaactaac        | 3000 |
| QY | 3271 | acactgaaacctgtcctcatgaggagacaagtgcagatcaggaatttgctctgttttcacc       | 3330 |
| Db | 3001 | acactgaaacctgtcctcatgaggagacaagtgcagatcaggaatttgctctgttttcacc       | 3060 |

| Db       | 4141  | ggagtaactactacgaaggagtaaaatctctctgtctacccagcatgatabtgaaaggattcttc | 4200 |
|----------|---|---|------|
| YY       | 4471  | acctcagcagctcaagatggcgcatcagtgatctctctttttcgaatgycaaagttaa        | 4530 |
| Db       | 4201  | acctcagcagatcaagaatgagccatcagtgatctctcttttcagaaatgycaaagttaa      | 4260 |
| YY       | 4531  | gtttttcaggaaatcaagaatccctctacacactgtgttgaaactctcagaccacagtta      | 4590 |
| Db       | 4261  | gtttttcaggaaatcaagaatccctctacacactgtgttgaaactctcagaccacagtta      | 4320 |
| YY       | 4591  | ctgactcgtctacacttcgaattcacccccagagttgggtgcacagatgcccctgagga       | 4650 |
| Db       | 4321  | ctgactcgtctacacttcgaattcacccccagagttgggtgcacagatgcccctgagga       | 4380 |
| YY       | 4651  | ggaggtctgggtgcgagggcacagagaccttactgactgactgagcaggttc              | 4699 |
| Db       | 4381  | ggaggtctgggtgcgagggcacagagaccttactgactgactgagcaggttc              | 4429 |
| RESULT 2 |   |   |      |
| AA       | AA73164   | standard; CDNA: 9354 BP.  |      |
| AC       | AA73164   |   |      |
| XX       | AA73164   |   |      |
| DT       | 08-APR-1998   | (first entry)   |      |
| XX       |   |   |      |
| DE       | CDNA encoding human B-domain deleted factor VIII.                         |   |      |
| XX       |   |   |      |
| KM       | Post-translational regulatory element; PRF; enhancer II; intronless gene; |   |      |
| KW       | surface antigen gene; cytoplasmic accumulation; targeted delivery;        |   |      |
| KV       | near consensus splice sequence; blood coagulation factor; factor VIII;    |   |      |
| KW       | factor IX; ss.  |   |      |
| XX       |   |   |      |
| OS       | Homo sapiens.   |   |      |
| XX       |   |   |      |
| FH       | Key   | Location/Qualifiers   |      |
| FT       | CDS   | 2965..7380  |      |
| FT       |   | /*tag= a  |      |
| FT       | misc_feature  | 5165..5174  |      |
| FT       |   | /*tag= b  |      |
| FT       | misc_feature  | 5695..5703  |      |
| FT       |   | /*tag= c  |      |
| FT       | misc_feature  | 6320..6328  |      |
| FT       |   | /*tag= d  |      |
| FT       | misc_feature  | 6595..5603  |      |
| FT       |   | /*tag= e  |      |
| FT       | misc_feature  | 7045..7053  |      |
| FT       |   | /*tag= f  |      |
| FT       | misc_feature  | 7143..7152  |      |
| FT       |   | /*tag= g  |      |
| FT       | misc_feature  | 3296..3312  |      |
| FT       |   | /*tag= h  |      |
| FT       | misc_feature  | 4798..4817  |      |
| FT       |   | /*tag= i  |      |
| FT       | misc_feature  | 5023..5045  |      |
| FT       |   | /*tag= j  |      |
| FT       | misc_feature  | 5333..5355  |      |
| FT       |   | /*tag= k  |      |
| FT       | misc_feature  | 5520..5538  |      |
| FT       |   | /*tag= l  |      |
| FT       |   | /*tag= m  |      |
| FT       |   | /*tag= n  |      |
| FT       |   | /*tag= o  |      |
| FT       |   | /*tag= p  |      |
| FT       |   | /*tag= q  |      |
| FT       |   | /*tag= r  |      |
| FT       |   | /*tag= s  |      |
| FT       |   | /*tag= t  |      |
| FT       |   | /*tag= u  |      |
| FT       |   | /*tag= v  |      |
| FT       |   | /*tag= w  |      |
| FT       |   | /*tag= x  |      |
| FT       |   | /*tag= y  |      |
| FT       |   | /*tag= z  |      |
| FT       |   | /*tag= 0  |      |
| FT       |   | /*tag= 1  |      |
| FT       |   | /*tag= 2  |      |
| FT       |   | /*tag= 3  |      |
| FT       |   | /*tag= 4  |      |
| FT       |   | /*tag= 5  |      |
| FT       |   | /*tag= 6  |      |
| FT       |   | /*tag= 7  |      |
| FT       |   | /*tag= 8  |      |
| FT       |   | /*tag= 9  |      |
| FT       |   | /*tag= A  |      |
| FT       |   | /*tag= B  |      |
| FT       |   | /*tag= C  |      |
| FT       |   | /*tag= D  |      |
| FT       |   | /*tag= E  |      |
| FT       |   | /*tag= F  |      |
| FT       |   | /*tag= G  |      |
| FT       |   | /*tag= H  |      |
| FT       |   | /*tag= I  |      |
| FT       |   | /*tag= J  |      |
| FT       |   | /*tag= K  |      |
| FT       |   | /*tag= L  |      |
| FT       |   | /*tag= M  |      |
| FT       |   | /*tag= N  |      |
| FT       |   | /*tag= O  |      |
| FT       |   | /*tag= P  |      |
| FT       |   | /*tag= Q  |      |
| FT       |   | /*tag= R  |      |
| FT       |   | /*tag= S  |      |
| FT       |   | /*tag= T  |      |
| FT       |   | /*tag= U  |      |
| FT       |   | /*tag= V  |      |
| FT       |   | /*tag= W  |      |
| FT       |   | /*tag= X  |      |
| FT       |   | /*tag= Y  |      |
| FT       |   | /*tag= Z  |      |
| FT       |   | /*tag= [  |      |
| FT       |   | /*tag= \  |      |
| FT       |   | /*tag= ]  |      |
| FT       |   | /*tag= ^  |      |
| FT       |   | /*tag= _  |      |
| FT       |   | /*tag= `  |      |
| FT       |   | /*tag= a  |      |
| FT       |   | /*tag= b  |      |
| FT       |   | /*tag= c  |      |
| FT       |   | /*tag= d  |      |
| FT       |   | /*tag= e  |      |
| FT       |   | /*tag= f  |      |
| FT       |   | /*tag= g  |      |
| FT       |   | /*tag= h  |      |
| FT       |   | /*tag= i  |      |
| FT       |   | /*tag= j  |      |
| FT       |   | /*tag= k  |      |
| FT       |   | /*tag= l  |      |
| FT       |   | /*  |      |

|    |   |                       |                      |
|----|---|-----------------------|----------------------|
| FT | misc_feature  | 5604..5632            | /tag= m              |
| FT |   | /note= "3'            | near consensus site" |
| FT | misc_feature  | 5717..5745            | /tag= n              |
| FT |   | /note= "3'            | near consensus site" |
| FT | misc_feature  | 6239..6258            | /tag= o              |
| FT |   | /note= "3'            | near consensus site" |
| FT | misc_feature  | 6658..6682            | /tag= p              |
| FT |   | /note= "3'            | near consensus site" |
| FT | misc_feature  | 7159..7176            | /tag= q              |
| FT |   | /note= "3'            | near consensus site" |
| FT | misc_feature  | 7196..7209            | /tag= r              |
| FT |   | /note= "3'            | near consensus site" |
| FT | misc_feature  | 7289..7315            | /tag= s              |
| FT |   | /note= "3'            | near consensus site" |
| FT | misc_feature  | 7411..7429            | /tag= t              |
| FT |   | /note= "3'            | near consensus site" |
| FT | misc_feature  | 7611..8197            | /tag= u              |
| FT |   | /note= "PRE sequence" |                      |
| XX | W09733994-A1.   |                       |                      |
| XX | 18-SEP-1997.  |                       |                      |
| XX | 10-MAR-1997:  | 97MO-US03561.         |                      |
| XX | 11-MAR-1996:  | 96US-0683839.         |                      |
| XX | (IMMU-) IMMUNE RESPONSE CORP.   |                       |                      |
| XX | Bidlingmaier S, III CR;   |                       |                      |
| XX | WPI: 1997-470874/43.  |                       |                      |
| XX | P-PSDB: AAW23414.   |                       |                      |
| PT | Vector for increased expression of intronless genes - comprises           |                       |                      |
| PT | intronless gene with at least one near consensus splice sequence, a       |                       |                      |
| PT | promoter and at least one viral cis-acting post-transcriptional           |                       |                      |
| PT | regulatory element  |                       |                      |
| XX | Example 1: Pages 21-31; 59pp: English.                                    |                       |                      |
| XX | The present sequence represents human B-domain deleted factor VIII        |                       |                      |
| CC | cDNA, and a post-translational regulatory element (PRE) of the            |                       |                      |
| CC | Hepatitis B virus, which is present 3' of the STOP codon for factor VIII. |                       |                      |
| CC | PRE sequences have been shown to function in cis to increase the          |                       |                      |
| CC | steady-state levels of surface gene transcripts by facilitating           |                       |                      |
| CC | cytoplasmic accumulation of these transcripts. The present sequence       |                       |                      |
| CC | is part of a novel vector, comprising an intronless gene containing       |                       |                      |
| CC | 1 or more near consensus splice sequences operably linked to a            |                       |                      |
| CC | promoter sequence so that the gene is transcribed in a cell.              |                       |                      |
| CC | Intronless gene transcripts which contain near consensus splice site      |                       |                      |
| CC | sequences are believed to get tied up in the nucleus of the cell where    |                       |                      |
| CC | splicing occurs, rather than being transported to the cytoplasm where     |                       |                      |
| CC | they can be translated into proteins. The PRE sequences are transcribed   |                       |                      |
| CC | along with the gene, causing export of the gene transcript from the       |                       |                      |
| CC | nucleus into the cytoplasm of the cell. The vector can be used            |                       |                      |
| CC | to increase the expression of an intronless gene containing at least one  |                       |                      |
| CC | near consensus splice sites, preferably cDNA encoding a blood coagulation |                       |                      |
| CC | factor, particularly Factor VIII or IX. The complex allows the targeted   |                       |                      |
| CC | delivery of the vector to a specific cell, e.g. hepatocytes when the      |                       |                      |
| CC | ligand is an asialoglycoprotein which binds the asialoglycoprotein        |                       |                      |
| CC | receptor present on their surface.  |                       |                      |
| XX | Sequence 9354 BP; 2506 A; 2239 C; 2161 G; 2448 T; 0 other;                |                       |                      |

|                       |                 |                |           |              |
|-----------------------|-----------------|----------------|-----------|--------------|
| Query Match           | 92.6%;          | Score 4411.6;  | DB 18;    | Length 9354; |
| Best Local Similarity | 99.7%;          | Pred. No. 0;   |           |              |
| Matches 4420;         | Conservative 0; | Mismatches 14; | Indels 0; | Gaps 0;      |

|    |      |  |      |
|----|------|--|------|
| OY | 266  | ccacacatgcaataatagagctctccaccgctctctctctgcttctgacattcgtctta          | 325  |
| Dp | 2960 | ccacacatgctttaatgagctctccaccgctctctctctctgcttctgacattcgtctta         | 3019 |
| OY | 326  | gtgcacacagaataatactaccctggcgagctggaactctcaatgggactataatgaagctg       | 385  |
| Dp | 3020 | gtgcacacagaataatactaccctggcgagctggaactctcaatgggactataatgaagctg       | 3079 |
| OY | 386  | atctcgtgtgagctgcctgttgagcgcaagaatttcctcccttagagctgcacaaatctttccat    | 445  |
| Dp | 3080 | atctcgtgtgagctgcctgttgagcgcaagaatttcctcccttagagctgcacaaatctttccat    | 3139 |
| OY | 446  | tcaaacacctgaatcgctgtatcaaaaagacctctgtttgtatgaattcaacgtttcaactttca    | 505  |
| Dp | 3140 | tcaaacacctgaatcgctgtatcaaaaagacctctgtttgtatgaattcaacgtttcaactttca    | 3199 |
| OY | 506  | acatcgcgttagcgcaagcgccacccttgatggtgcctgtatggtcttaccatccacgcgttagg    | 565  |
| Dp | 3200 | acatcgcgttagcgcaagcgccacccttgatggtgcctgtatggtcttaccatccacgcgttagg    | 3259 |
| OY | 566  | tttatgataaagctggtgcattacacttaagaataatgagcttccacatcgtctacgtttcatgt    | 625  |
| Dp | 3260 | tttatgataaagctggtgcattacacttaagaataatgagcttccacatcgtctacgtttcatgt    | 3319 |
| OY | 626  | ctgttgtgtatctactctacgtgaaagctctctgagggagctgaatgatgatcagaacagctc      | 685  |
| Dp | 3320 | ctgttgtgtatctactctacgtgaaagctctctgagggagctgaatgatgatcagaacagctc      | 3379 |
| OY | 686  | aaaggggaaagaagaatgaataagctctccctggcggtgggaagccatataatgtctcggaag      | 745  |
| Dp | 3380 | aaaggggaaagaagaatgaataagctctccctggcggtgggaagccatataatgtctcggaag      | 3439 |
| OY | 746  | tctctgaaagagaatggtcccaatggtcctctgaccacgtgtgccttaccatcatatctt         | 805  |
| Dp | 3440 | tctctgaaagagaatggtcccaatggtcctctgaccacgtgtgccttaccatcatatctt         | 3499 |
| OY | 806  | ctcatgtgtgacccctgtgaaagaagcttgaattcaagcgctcatattgggggcccctaatgatgata | 865  |
| Dp | 3500 | ctcatgtgtgacccctgtgaaagaagcttgaattcaagcgctcatattgggggcccctaatgatgata | 3559 |
| OY | 866  | gagagaaggagatctgtgccaagagaataagacacagaccctgtgcacaaattatactacttttg    | 925  |
| Dp | 3560 | gagagaaggagatctgtgccaagagaataagacacagaccctgtgcacaaattatactacttttg    | 3619 |
| OY | 926  | ctgtattttgatatgaagggaagaagtgtgacctacacgaataaagaatcctctgtatgcagata    | 985  |
| Dp | 3620 | ctgtattttgatatgaagggaagaagtgtgacctacacgaataaagaatcctctgtatgcagata    | 3679 |
| OY | 986  | gggagtgctgatactgtctgggcccctggcccttaaaatgcacacagctcaatggttatgttaaaca  | 1045 |
| Dp | 3680 | gggagtgctgatactgtctgggcccctggcccttaaaatgcacacagctcaatggttatgttaaaca  | 3739 |
| OY | 1046 | ggtctctgcacaggtctgattatgtgatgccaaggaataatcgatctatgtgcattgtgatgtgaat  | 1105 |
| Dp | 3740 | ggtctctgcacaggtctgattatgtgatgccaaggaataatcgatctatgtgcattgtgtgaat     | 3799 |
| OY | 1106 | tgggacacacatcctctgtaagtgtgcactcaatatctctctgaaaggtcacaacattctgttgagga | 1165 |
| Dp | 3800 | tgggacacacatcctctgtaagtgtgcactcaatatctctctgaaaggtcacaacattctgttgagga | 3859 |
| OY | 1166 | accatgcgcaggggtctctctgtgaataatctgcgcacataaacttctctatgtgtcaaacactc    | 1225 |
| Dp | 3860 | accatgcgcaggggtctctctgtgaataatctgcgcacataaacttctctatgtgtcaaacactc    | 3919 |
| OY | 1226 | tgatgatcctgtgacaagttctactactgttttgatcatatctctccacacacatgatgtga       | 1285 |
| Dp | 3920 | tgatgatcctgtgacaagttctactactgttttgatcatatctctccacacacatgatgtga       | 3979 |

|    |      |  |      |
|----|------|--|------|
| QY | 1266 | tggaagcttatgttcaaaagtacagctgttccaagagaaaccccaactaagatgaaata        | 1345 |
| DB | 3980 | tggaaagctatgttcaaaatlagcaagctgttccaagagaaaccccaactaagatgaaata      | 4039 |
| QY | 1346 | atgaagaagcggaaagacatgtatgtatgtcttaacgacttcctgaaatgtatgtacagt       | 1405 |
| DB | 4040 | atgaagaagcggaaagacatgtatgtatgtcttaacgacttcctgaaatgtatgtacagt       | 4099 |
| QY | 1406 | tgtatgtatgacaactctctccttcttataccaattcgctcagttgtgccaagaactccta      | 1465 |
| DB | 4100 | tgtatgtatgacaactctctccttcttataccaattcgctcagttgtgccaagaagactccta    | 4159 |
| QY | 1466 | aaacttgggtacaattacaattgtctgttgaagggagagacttggagactgtcttccttaagcc   | 1525 |
| DB | 4160 | aaacttgggtacaattacaattgtctgttgaagggagagacttggagactgtcttccttaagcc   | 4219 |
| QY | 1536 | tgcgccccgagtacagaagttataaaagtccaatatgttgaacaatgtgccttcagcggatgtg   | 1585 |
| DB | 4220 | tgcgccccgagtacagaagttataaaagtccaatatgttgaacaatgtgccttcagcggatgtg   | 4279 |
| QY | 1586 | gttagaagtacaaaaaagttccgatttatgtgcatacaacagatgaaaaccttaagactgtgt    | 1645 |
| DB | 4280 | gttagaagtacaaaaaagttccgatttatgtgcatacaacagatgaaaaccttaagactgtgt    | 4339 |
| QY | 1646 | aagcatttcagatgaatccaagaactcttggagaccttcttatatgtgggaagtgtgagaca     | 1705 |
| DB | 4340 | aagcatttcagatgaatccaagaactcttggagaccttcttatatgtgggaagtgtgagaca     | 4399 |
| QY | 1706 | caactgtatataatlttaagaataccaagaagacagacataatacattcacctccagga        | 1765 |
| DB | 4400 | caactgtatataatlttaagaataccaagaagacagacataatacattcacctccagga        | 4459 |
| QY | 1766 | tcaactgagtccgcttcttctgttatccaagagatataccaagaagtggtaaacatttgaag     | 1825 |
| DB | 4460 | tcaactgagtccgcttcttctgttatccaagagatataccaagaagtggtaaacatttgaag     | 4519 |
| QY | 1836 | atttccaaattctgcagagagaaatatattcaaatataatgtgacagtgtactgtatgaagatg   | 1885 |
| DB | 4520 | atttccaaattctgcagagagaaatatattcaaatataatgtgacagtgtactgtatgaagatg   | 4579 |
| QY | 1886 | ggccaactcaaatccagatccctcctgtgtccgcagaccgctatactctagtttcgttaatatg   | 1945 |
| DB | 4580 | ggccaactcaaatccagatccctcctgtgtccgcagaccgctatactctagtttcgttaatatg   | 4639 |
| QY | 1946 | agaagatcatagcttcagagactcattgtgcctctcctcatctgtctatacaagaactctgag    | 2005 |
| DB | 4640 | agaagatcatagcttcagagactcattgtgcctctcctcatctgtctatacaagaactctgag    | 4699 |
| QY | 2006 | atcaaaagsgaaaccagataatgttcagaacaagaagatgtcaatccgtgtttctgtatttg     | 2065 |
| DB | 4700 | atcaaaagsgaaaccagataatgttcagaacaagaagatgtcaatccgtgtttctgtatttg     | 4759 |
| QY | 2066 | atgaaaacccgaagctgtgtaaccttcacagaagaatatatacaagctcttctcccaatccagctg | 2125 |
| DB | 4760 | atgaaaacccgaagctgtgtaaccttcacagaagaatatatacaagctcttctcccaatccagctg | 4819 |
| QY | 2126 | gagtcgaactttagagatccagatccagttccaaagctcccaaatcatgtcacagacataatgct  | 2185 |
| DB | 4820 | gagtcgaactttagagatccagatccagttccaaagctcccaaatcatgtcacagacataatgct  | 4879 |
| QY | 2186 | atgttttttgataagtttgcagttgttgcagttgttgcatatgagatgtgcatactgtgtacatc  | 2245 |
| DB | 4880 | atgttttttgataagtttgcagttgttgcagttgttgcatatgagatgtgcatactgtgtacatc  | 4939 |
| QY | 2246 | taagatctggagacaaagctgactctcttctgtctctctctctctctgagataatccttcaac    | 2305 |
| DB | 4940 | taagatctggagacaaagctgactctcttctgtctctctctctctctgagataatccttcaac    | 4999 |
| QY | 2306 | acaaaaatgtctatagaacacaactacacctattccattctccaggaagaatgtcttca        | 2365 |
| DB | 5000 | acaaaaatgtctatagaacacaactacacctattccattctccaggaagaatgtcttca        | 5059 |



QY 2366 tgcgatggaaccacgagctatgagtlctgagggtgcacaactcagacttcggaaca 2425  
|||||  
Db 5060 tgcgatggaaccacgagctatgagtlctgagggtgcacaactcagacttcggaaca 5119  
QY 2426 gagcagagaccgcttaactgaagtttctagtgtgacaaagaacctgggtattacg 2485  
|||||  
Db 5120 gagcagagaccgcttaactgaagtttctagtgtgacaaagaacctgggtattacg 5179  
QY 2486 aggcacgttaagaagatcttccagactatgctgagtaaaacaatccatctgaacca 2545  
|||||  
Db 5180 aggcacgttaagaagatcttccagactatgctgagtaaaacaatccatctgaacca 5239  
QY 2546 gaagcttctccagaatccaagaacccctagcactagaagcaaaagaatttaatgccc 2605  
|||||  
Db 5240 gaagcttctccagaatccaagaacccctagcactagaagcaaaagaatttaatgccc 5299  
QY 2606 caccagctctgaaacgcgcatcaacgggaaataactcgtactctcttcagttacgaag 2665  
|||||  
Db 5300 caccagctctgaaacgcgcatcaacgggaaataactcgtactctcttcagttacgaag 5359  
QY 2666 aggaatctgactatgatatccatatcagttgaaatgaagaaggaatttgcacatt 2725  
|||||  
Db 5360 aggaatctgactatgatatccatatcagttgaaatgaagaaggaatttgcacatt 5419  
QY 2726 atgattgagatgaaatcagaagccccgcagcttcaaaagaaacacgacatactta 2785  
|||||  
Db 5420 atgattgagatgaaatcagaagccccgcagcttcaaaagaaacacgacatactta 5479  
QY 2786 ttgctgagttgagaaggtctgagattatggaatgagttccccaatgtctcaagaa 2845  
|||||  
Db 5480 ttgctgagttgagaaggtctgagattatggaatgagttccccaatgtctcaagaa 5539  
QY 2846 acagaggtcagagatgagcagtgccctcagttccaagaaggttgttccaagatttaccg 2905  
|||||  
Db 5540 acagaggtcagagatgagcagtgccctcagttccaagaaggttgttccaagatttaccg 5599  
QY 2906 atggtccttactcagccctataacggtggaagaactaaatgaaacttggagctctgg 2965  
|||||  
Db 5600 atggtccttactcagccctataacggtggaagaactaaatgaaacttggagctctgg 5659  
QY 2966 ggccattataagggcagaattgaaatataatcattgtaacttccgaatccagct 3025  
|||||  
Db 5660 ggccattataagggcagaattgaaatataatcattgtaacttccgaatccagct 5719  
QY 3026 ctgcctccatctcctctatctagccttacttcttaaggaagaatcagaagcagaag 3085  
|||||  
Db 5720 ctgcctccatctcctctatctagccttacttcttaaggaagaatcagaagcagaag 5779  
QY 3086 cagaactagaanaaacttgcagccttaagccttaatgaaaccaaacttacttggaaatg 3145  
|||||  
Db 5780 cagaactagaanaaacttgcagccttaagccttaatgaaaccaaacttacttggaaatg 5839  
QY 3146 aacatcatatggcaccactaaagatgagttgactgcaagcctgggcttaattctctg 3205  
|||||  
Db 5840 aacatcatatggcaccactaaagatgagttgactgcaagcctgggcttaattctctg 5899  
QY 3206 atgttgaccttgaaanaaagatgtgcactcagcctgatttgagccctctctgctgcaca 3265  
5900 atgttgaccttgaaanaaagatgtgcactcagcctgatttgagccctctctgctgcaca 5959  
QY 3266 ctacaacaactgaaacccgtctcaatgggagacaagtgcacagatccaggaatttgcctg 3325  
5960 ctacaacaactgaaacccgtctcaatgggagacaagtgcacagatccaggaatttgcctg 6019  
Db 3326 tcaacatcttggatgagaccaaaagctgtacttcaactgtaaaatggaagaaactca 3385  
6020 tcaacatcttggatgagaccaaaagctgtacttcaactgtaaaatggaagaaactca 6079  
QY 3386 gggtcccttgcataatccagaatggaagatcccaactttaaagagatatcgtctccag 3445  
6080 gggtcccttgcataatccagaatggaagatcccaactttaaagagatatcgtctccag 6139  
QY 3446 caatcaatggtctacataatgatacactaactggtctagtaatggctcagatcaagaa 3505

Db 6140 caatcaatggtctacataatgatacactaactggtctagtaatggtcagatcaagaa 6199  
QY 3506 ttccagtttatctgtctcagcagtgaggcagcaatggaacatcatcttatcttaccgtg 3565  
6200 ttccagtttatctgtctcagcagtgaggcagcaatggaacatcatcttatcttaccgtg 6259  
QY 3566 gacatgtctcactgtacgaanaaagaggtataaaatgacactgtacatctcactc 3625  
6260 gacatgtgtctcactgtacgaanaaagaggtataaaatgacactgtacatctcactc 6319  
Db 6320 caggtgttttgagacagtggaatglttaccatccaagctggaatttgggggtggag 6379  
QY 3686 gcttattggcagagatctacatgtctggatgagacacctttctgtgtgacagcata 3745  
6380 gcttattggcagagatctacatgtctggatgagacacctttctgtgtgacagcata 6439  
QY 3746 agtgcagaatcccttgggaatggtctctggaacacatlagatlttccagatccagct 3805  
6440 agtgcagaatcccttgggaatggtctctggaacacatlagatlttccagatccagct 6499  
QY 3806 caggaacaataggaagttgggcccccaagctggcagagatctatctcggatccatca 3865  
6500 caggaacaataggaagttgggcccccaagctggcagagatctatctcggatccatca 6559  
QY 3866 atgctctgagacccaagagacccttctctgatatcaagtgatcgtgtgacccaatga 3925  
6560 atgctctgagacccaagagacccttctctgatatcaagtgatcgtgtgacccaatga 6619  
Db 3926 ttattcaaggaatcaagaaccaggtgtcccgltcagaagttctccagcctctacatctc 3985  
6620 ttattcaaggaatcaagaaccaggtgtcccgltcagaagttctccagcctctacatctc 6679  
QY 3986 agttatcatcatgtatgcttctgattggaagaggtgagacattatccgaggaattcca 4045  
6680 agttatcatcatgtatgcttctgattggaagaggtgagacattatccgaggaattcca 6739  
Db 4046 ctggaaccttaatgtctctcttctggaatgtggttcaatctcgggatataaacaatat 4105  
6740 ctggaaccttaatgtctctcttctggaatgtggttcaatctcgggatataaacaatat 6799  
QY 4106 ttaacctccaattatgtctcgaataacacgtcttgacaccaacccatlatagcatctgca 4165  
6800 ttaacctccaattatgtctcgaataacacgtcttgacaccaacccatlatagcatctgca 6859  
QY 4166 gcaactcttcgcatgagttgattgggtctgtgatttaaatgttgcagcatgccaattggaa 4225  
6860 gcaactcttcgcatgagttgattgggtctgtgatttaaatgttgcagcatgccaattggaa 6919  
Db 4226 tggagagtcaagcaatcagatgacagatctactgcttcaacttcaacttcaacttga 4285  
6920 tggagagtcaagcaatcagatgacagatctactgcttcaacttcaacttcaacttga 6979  
QY 4286 ttgccaactgtgtctctcaaaagctgcagactccaactccaaggagagtaatgtcctgga 4345  
6980 ttgccaactgtgtctctcaaaagctgcagactccaactccaaggagagtaatgtcctgga 7039  
QY 4346 gacctcaggtgaatataatccaagaaggtgtgcgaagtgagatccagaagaacatgaaag 4405  
7040 gacctcaggtgaatataatccaagaaggtgtgcgaagtgagatccagaagaacatgaaag 7099  
Db 4406 tcaacagagtaactactcagggagtaaaatctcgtcttaccagcatgtatgtgaagagt 4465  
7100 tcaacagagtaactactcagggagtaaaatctcgtcttaccagcatgtatgtgaagagt 7159  
QY 4466 tctcatatccagcagttcaagaatggtccatcagtgagactcttlttccgaatggtcaag 4525  
7160 tctcatatccagcagttcaagaatggtccatcagtgagactcttlttccgaatggtcaag 7219  
QY 4526 taaagttttccagggaatacagaactcctccaacactgtgtgaaactctctagaccac 4585

Db 7220 taaggattttcaggaaatcaagactcttcacacctgtgtgaactctctagaccac 7279  
QY 4586 cgttactgactcgtactcttgaattcaacccccagattgtgtgacacagattccctga 4645  
Db 7280 cgttactgactcgtactcttgaattcaacccccagattgtgtgacacagattccctga 7339  
QY 4646 ggaagagaggtcttggtcgtcgaagacagacacctactctactcgaagcgaattc 4699  
Db 7340 ggaagagaggtcttggtcgtcgaagacagacacctactctactcgaaggtgagcactgc 7393

RESULT 3  
AA076016  
ID AA076016 standard; cDNA; 4629 BP.  
XX  
AC AA076016;  
XX  
DT 20-JUL-1995 (first entry)  
XX  
DE B-domain deleted Factor-VIII.  
XX  
KM Factor-VIII; blood-clotting; hemophilia A; gene therapy;  
KM adenovirus; vector; ss.  
XX  
OS Homo sapiens.  
XX  
PN M09429471-A.  
XX  
PD 22-DEC-1994.  
XX  
PE 13-APR-1994; 94MO-US04075.  
XX  
PR 10-JUN-1993; 93US-0074920.  
PR 25-MAR-1994; 94US-0218335.  
XX  
PA (GENE-) GENETIC THERAPY INC.  
XX  
PI Connelly S, Kaleko M, Smith T;  
XX  
DR WPI; 1995-036495/05.  
DR P-PSDB; AAR67709.  
XX  
PT New adenoviral vectors for treatment of haemophilia - contg. a  
PT DNA sequence encoding a clotting factor, partic. Factor VIII or  
PT Factor IX  
XX  
PS Disclosure; Fig. 17A-17C; 116pp; English.  
XX  
CC Human Factor-VIII cDNA, from which the B domain had been deleted, was  
CC used to construct recombinant adenovirus vectors that produced  
CC therapeutic levels of the clotting factor when administered to an  
CC animal host, potentially providing hemophilia A gene therapy.  
XX  
SQ Sequence 4629 BP; 1318 A; 1051 C; 1018 G; 1242 T; 0 other;

Query Match 91.4%; Score 4355.4; DB 16; Length 4629;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 4383; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 271 atggaataagactctcaccctgctcttcttctgtcctttgtgacatctgtttagtgc 330  
Db 1 atggaataagactctcaccctgctcttcttctgtcctttgtgacatctgtttagtgc 60  
QY 331 accagaagatactacctggtgtgcagtggaactgtcatggaactatgtcaaaagtatctc 390  
Db 61 accagaagatactacctggtgtgcagtggaactgtcatggaactatgtcaaaagtatctc 120  
QY 391 ggtgagctgctgtgtgacgcaagaattccctcctagagtgccaaaatctttccattcaac 450  
Db 121 ggtgagctgctgtgtgacgcaagaattccctcctagagtgccaaaatctttccattcaac 180  
QY 451 acccagtcgtgtacaaaagactctgtttgttagaatcaagtcacacttccaacatc 510

Db 181 acccagtcgtgtacaaaagactctgtttgttagaatcaagtcacacttccaacatc 240  
QY 511 gctaagccaagccaacctgtgagtgtctgtaagttccatccacagctgaggttat 570  
Db 241 gctaagccaagccaacctgtgagtgtctgtaagttccatccacagctgaggttat 300  
QY 571 gatacagtggtcattacacttaagaacatggtctccatccctgtcagttcattagtggtt 630  
Db 301 gatacagtggtcattacacttaagaacatggtctccatccctgtcagttcattagtggtc 360  
QY 631 ggttatccctacttgaaagctctcgaaggagctgaatagtatgtacagaccagcaagg 690  
Db 361 ggttatccctacttgaaagctctcgaaggagctgaatagtatgtacagaccagcaagg 420  
QY 691 gagaagaagaatgtataaagctctccctgtgtgaaagccaataatgtctgcaagctctg 750  
Db 421 gagaagaagaatgtataaagctctccctgtgtgaaagccaataatgtctgcaagctctg 480  
QY 751 aaagaagaatgttccaatgtgctctgacccactgtgacctactactatcttctcat 810  
Db 481 aaagaagaatgttccaatgtgctctgacccactgtgacctactactatcttctcat 540  
QY 811 gtggaacctgtgtaaaagacttgaaattcaagcctcatgtgagccctactagtatagaagaa 870  
Db 541 gtggaacctgtgtaaaagacttgaaattcaagcctcatgtgagccctactagtatagaagaa 600  
QY 871 gggagctgtgccaaggaagaacacagacacttgacaaattatactacttttgcgtta 930  
Db 601 gggagctgtgccaaggaagaacacagacacttgacaaattatactacttttgcgtta 660  
QY 931 ttgatgaaggaagaaagttgtgactctcgaagaacaaagaactccttgatgcaagataggat 990  
Db 661 ttgatgaaggaagaaagttgtgactctcgaagaacaaagaactccttgatgcaagataggat 720  
QY 991 gctgcatctgtcgcggcctcgtcgtcaaaatgtgcacacagtcgaattgttgaataacagctc 1050  
Db 721 gctgcatctgtcgcggcctcgtcgtcgtcaaaatgtgcacacagtcgaattgttgaataacagctc 780  
QY 1051 ctgcacagctctgattgtagtgcacacagaagaatcagttctatgtgcatgtgaaatgggc 1110  
Db 781 ctgcacagctctgattgtagtgcacacagaagaatcagttctatgtgcatgtgaaatgggc 840  
QY 1111 accactctgaagtgacactaatacttctcgaagtgcaacattctgtgtaaggaacat 1170  
Db 841 accactctgaagtgacactaatacttctcgaagtgcaacattctgtgtaaggaacat 900  
QY 1171 cgcagagcgtctcgtgaaatctgcgcaataacttctcctactgtcgaacactctgtatg 1230  
Db 901 cgcagagcgtctcgtgaaatctgcgcaataacttctcctactgtcgaacactctgtatg 960  
QY 1231 gacctgtgacagtttctactcgttttgttcatactctctccacccaacatgcatgtgaa 1290  
Db 961 gacctgtgacagtttctactcgttttgttcatactctctccacccaacatgcatgtgaa 1020  
QY 1291 gcttatgtcaaaagtacagctgttccagaggaaccccacacagaaatgaaataatgaa 1350  
Db 1021 gcttatgtcaaaagtacagctgttccagaggaaccccacacagaaatgaaataatgaa 1080  
QY 1351 gaagcgaagactatgatagatcttactatctcgtcgaagatggaatgtgtcaggtttat 1410  
Db 1081 gaagcgaagactatgatagatcttactatctcgtcgaagatggaatgtgtcaggtttat 1140  
QY 1411 gatgacaactctctctcttataccaaattgtcgtcagttgccaagaacatctctaaact 1470  
Db 1141 gatgacaactctctctcttataccaaattgtcgtcagttgccaagaacatctctaaact 1200  
QY 1471 tgggtacattatcatctgtcgtcgaagagagagcgtggagactatgttccttactgtctgcgc 1530  
Db 1201 tgggtacattatcatctgtcgtcgaagagagagcgtggagactatgttccttactgtctgcgc 1260  
QY 1531 cccgatatcaagaagttaataaagtcgaatatgtgacaatagtcctcagcagatgtgttagg 1590

Dh 1261 ccgacgacgaagatlaataaagltcaatalttgaaacaaatggccctcagcggatgg 1320  
Qy 1591 aagfacaaaaaagtcocgattatgtgacacacaga tgaaccccttaagaactgttaagct 1650  
Dh 1321 aagfacaaaaaagtcocgattatgtgacacacaga tgaaccccttaagaactgttaagct 1380  
Qy 1651 attcagcatgaaatcaagaatcttgggaccttacttta tggggaagtggagacacactg 1710  
Dh 1381 attcagcatgaaatcaagaatcttgggaccttacttta tggggaagtggagacacactg 1440  
Qy 1711 ttgattatattgaagaatcaagaagcagaccata taaatctaccctcaggaatccct 1770  
Dh 1441 ttgattatattgaagaatcaagaagcagaccata taaatctaccctcaggaatccct 1500  
Qy 1771 gatctccgtcccttgcattcaagaagatatacaca aagggtgtaaaacatttgaagaattt 1830  
Dh 1501 gatctccgtcccttgcattcaagaagatatacaca aagggtgtaaaacatttgaagaattt 1560  
Qy 1831 ccaattctgcagaggaataattcaaatataatgga cagtgcgtgtgaagaatgggcca 1890  
Dh 1561 ccaattctgcagaggaataattcaaatataatgga cagtgcgtgtgaagaatgggcca 1620  
Qy 1891 actaaatcagaatccctcgtgctgaacccgctatta ctactagtttcgttaataatgaga 1950  
Dh 1621 actaaatcagaatccctcgtgctgaacccgctatta ctactagtttcgttaataatgaga 1680  
Qy 1951 gatctacgttcaggaacatctgcgcctcctcatal ctgtctcaaaaagaatctgtaacaa 2010  
Dh 1681 gatctacgttcaggaacatctgcgcctcctcatal ctgtctcaaaaagaatctgtaacaa 1740  
Qy 2011 agaggaacacagataatgltcagacagaagaatg tcatccgtgttctgtatattgag 2070  
Dh 1741 agaggaacacagataatgltcagacagaagaatg tcatccgtgttctgtatattgag 1800  
Qy 2071 aacccgaagctggtatccctcacagagaataata caaagcttctccccaatcccgctgagtg 2130  
Dh 1801 aacccgaagctggtatccctcacagagaataata caaagcttctccccaatcccgctgagtg 1860  
Qy 2131 cagcttagagatccagaggttccaagcctcacaac aatcagacagatcaatgtgcattt 2190  
Dh 1861 cagcttagagatccagaggttccaagcctcacaac aatcagacagatcaatgtgcattt 1920  
Qy 2191 ttgatatgttgcagctgtgcagttgttgcata gtaggtgcacactggtlaactaaagc 2250  
Dh 1921 ttgatatgttgcagctgtgcagttgttgcata gtaggtgcacactggtlaactaaagc 1980  
Qy 2251 attgggaacacagactggtcttctcttctgtctt ctctctgtgatatcaacttcaaaacaa 2310  
Dh 1981 attgggaacacagactggtcttctcttctgtctt ctctctgtgatatcaacttcaaaacaa 2040  
Qy 2311 atgtctatagaaacacatacctaattcccatctcag agagaactgtcttaactg 2370  
Dh 2041 atgtctatagaaacacatacctaattcccatctcag agagaactgtcttaactg 2100  
Qy 2371 atggaacaaacccaggtctcattggaatctctgg ggtgcacaactcagacttccggaacagagc 2430  
Dh 2101 atggaacaaacccaggtctcattggaatctctgg ggtgcacaactcagacttccggaacagagc 2160  
Qy 2431 atgaaccccttactaggaaggttctcagtgttga cagaagaactgtgatatcaagagac 2490  
Dh 2161 atgaaccccttactaggaaggttctcagtgttga cagaagaactgtgatatcaagagac 2220  
Qy 2491 agttatagaatatttcaagcaacttgcgtgataa aacaaatgcatctgaaccaaagagc 2550  
Dh 2221 agttatagaatatttcaagcaacttgcgtgataa aacaaatgcatctgaaccaaagagc 2280  
Qy 2551 ttctccacagaatltcaagaacacctagcaactag gcaaaaagaatattaatgccaacccacaa 2610  
Dh 2281 ttctccacagaatltcaagaacacctagcaactag gcaaaaagaatattaatgccaacccacaa 2340  
Qy 2611 gtcttgaagaagcgaatcaaggggaataactgtta ctactcttgcgtacgaatgaagaa 2670  
Dh 2341 gtcttgaagaagcgaatcaaggggaataactgtta ctactcttgcgtacgaatgaagaa 2400

Qy 2671 attgacatgatatcatcatactcaglttgaaatgaa gaaagagaatttvgacatttatg 2730  
Dh 2401 attgacatgatatcatcatactcaglttgaaatgaa gaaagagaatttvgacatttatg 2460  
Qy 2731 gaggatgaaatcagagcccccagagcttccaanaa gaacagacataatttatgct 2790  
Dh 2461 gaggatgaaatcagagcccccagagcttccaanaa gaacagacataatttatgct 2520  
Qy 2791 gcaagtggagagctcttgagattatggaatgagtg agctcccaactgtttcaagaagagc 2850  
Dh 2521 gcaagtggagagctcttgagattatggaatgagtg agctcccaactgtttcaagaagagc 2580  
Qy 2851 gtcagagtgagcagtgctccctcagttcaagaaga gttgtttccaggaatttactgagc 2910  
Dh 2581 gtcagagtgagcagtgctccctcagttcaagaaga gttgtttccaggaatttactgagc 2640  
Qy 2911 tcccttactcagcccttatccogtggagaactaaat gaaactttgggaccccttgggcca 2970  
Dh 2641 tcccttactcagcccttatccogtggagaactaaat gaaactttgggaccccttgggcca 2700  
Qy 2971 tatataagagcagaagttggaagataataatcag tttgatacttcagagaatccagccctcgt 3030  
Dh 2701 tatataagagcagaagttggaagataataatcag tttgatacttcagagaatccagccctcgt 2760  
Qy 3031 cccctatccctctatctcagccttattctta tgaagaagatcagagcagaagca 3090  
Dh 2761 cccctatccctctatctcagccttattctta tgaagaagatcagagcagaagca 2820  
Qy 3091 cctagaanaaaccttgcataagcctaaatgaaccaa aactactcttggaaagtgcacat 3150  
Dh 2821 cctagaanaaaccttgcataagcctaaatgaaccaa aactactcttggaaagtgcacat 2880  
Qy 3151 catatggacaccacataagaatgagtttgaactgc aagaagccttgggttatctctgagtt 3210  
Dh 2881 catatggacaccacataagaatgagtttgaactgc aagaagccttgggttatctctgagtt 2940  
Qy 3211 gaacctggaanaaagatgltgcactcagccctga tttggaacccctctcgtgtccacactaac 3270  
Dh 2941 gaacctggaanaaagatgltgcactcagccctga tttggaacccctctcgtgtccacactaac 3000  
Qy 3271 acactgaaacccctgtcctcatgtggagagaaga gttgacagttatgtctgttttacc 3330  
Dh 3001 acactgaaacccctgtcctcatgtggagagaaga gttgacagttatgtctgttttacc 3060  
Qy 3331 atctttgaatggaaccaaagcctgtgacttcaact gaataatggaagaagaactgcaggtc 3390  
Dh 3061 atctttgaatggaaccaaagcctgtgacttcaact gaataatggaagaagaactgcaggtc 3120  
Qy 3391 cccctgcaatatccagatggaagatcccaactttaa gaagaaattatcgcttccatgcaatc 3450  
Dh 3121 cccctgcaatatccagatggaagatcccaactttaa gaagaaattatcgcttccatgcaatc 3180  
Qy 3451 aatgctcatcaatgagatgatacactacgttgctta gtaatggtctcagatccaagaattcga 3510  
Dh 3181 aatgctcatcaatgagatgatacactacgttgctta gtaatggtctcagatccaagaattcga 3240  
Qy 3511 tggatactgctcagcatgagtgagcagaatgaanaa acatccatctatcaatccaatgagac 3570  
Dh 3241 tggatactgctcagcatgagtgagcagaatgaanaa acatccatctatcaatccaatgagac 3300  
Qy 3571 gtgtcactgtacgaaaaaagagagatataaaatg gacatgtatacctatccacagt 3630  
Dh 3301 gtgtcactgtacgaaaaaagagagatataaaatg gacatgtatacctatccacagt 3360  
Qy 3631 gtttttgaacagatgagaaatgttaccatccaagc tctggaatttggcgggtggaatgcctt 3690  
Dh 3361 gtttttgaacagatgagaaatgttaccatccaagc tctggaatttggcgggtggaatgcctt 3420  
Qy 3691 attggcagacatctaacatgctcggagatggaacac ttttctcgtgtgtacgaataaagtgt 3750  
Dh 3421 attggcagacatctaacatgctcggagatggaacac ttttctcgtgtgtacgaataaagtgt 3480

|    |      |  |      |
|----|------|--|------|
| OY | 3751 | caagctcccccgggaagtgctcttcggaacattagaatttcaatataagccttcagga         | 3811 |
| Db | 3481 | caagctcccccgggaagtgctcttcggaacattagaagatttcagatttaacgcttcagga      | 3540 |
| OY | 3811 | caatactgacagctgggcccccaaaagccgagccagacttcatctccgga tcaacaatgcc     | 3870 |
| Db | 3541 | caatactgacagctgggcccccaaaagccgagccagacttcatctccgga tcaacaagcc      | 3600 |
| OY | 3871 | ttggagcccaagagagcccttcttcgtatcaaggttgatctgtgtggcaacaagtatat        | 3930 |
| Db | 3601 | ttggagcccaagagagcccttcttcgtatcaaggttgatctgtgtggcaacaagtatat        | 3660 |
| OY | 3931 | caagcgatccaagccccagaggctgcgcgtccaaggcttcacagccttaattctccagtt       | 3990 |
| Db | 3661 | caagcgatccaagccccagaggctgcgcgtccaaggcttcacagccttaattctccagtt       | 3720 |
| OY | 3991 | atcatactgtatagctctctgtgtggaagaagctggcagacttaccggaatctccactgga      | 4050 |
| Db | 3721 | atcatactgtatagctctctgtgtggaagaagctggcagacttaccggaatctccactgga      | 3780 |
| OY | 4051 | acccttaatgctctcttcttggcaatgtgga tcaactcgtggataaaacacaatattttaac    | 4110 |
| Db | 3781 | acccttaatgctctcttcttggcaatgtgga tcaactcgtggataaaacacaatattttaac    | 3840 |
| OY | 4111 | ccctccaattatctctcgtatatacatccgttttgacccaactcatatagaatctcgacgaact   | 4170 |
| Db | 3841 | ccctccaattatctctcgtatatacatccgttttgacccaactcatatagaatctcgacgaact   | 3900 |
| OY | 4171 | cttcgcgaatgagcttgatctggcctgtgatttaaatagtttcagatgacatcttgggaatgag   | 4230 |
| Db | 3901 | cttcgcgaatgagcttgatctggcctgtgatttaaatagtttcagatgacatcttgggaatgag   | 3960 |
| OY | 4231 | agtaaaagcaatatacagaatgcacagaaattactgtcttcaactcatatcaacaatagtlttcgc | 4290 |
| Db | 3961 | agtaaaagcaatatacagaatgcacagaaattactgtcttcaactcatatcaacaatagtlttcgc | 4020 |
| OY | 4291 | accctggtctctctccaagaagctcgacttcaacccccaaggagaggaatgctctggaagact    | 4350 |
| Db | 4021 | accctggtctctctccaagaagctcgacttcaacccccaaggagaggaatgctctggaagact    | 4080 |
| OY | 4351 | caggtgataataccaagaagctgacgcgaagtggacttccaagaagacaatgaagtcaca       | 4410 |
| Db | 4081 | caggtgataataccaagaagctgacgcgaagtggacttccaagaagacaatgaagtcaca       | 4140 |
| OY | 4411 | ggagtaactactcaagggagtaaaatctctgtcttaccagaatgtatgtgaaggttctctc      | 4470 |
| Db | 4141 | ggagtaactactcaagggagtaaaatctctgtcttaccagaatgtatgtgaaggttctctc      | 4200 |
| OY | 4471 | atctccagcagatcaagaatgtgccaatcagctgagactccttttccaagaatgccaagaatgaag | 4530 |
| Db | 4201 | atctccagcagatcaagaatgtgccaatcagctgagactccttttccaagaatgccaagaatgaag | 4260 |
| OY | 4531 | gtttttcaaggaataccaagaactccttcaacaacctgtggagaaactcctagaccacacgtta   | 4590 |
| Db | 4261 | gtttttcaaggaataccaagaactccttcaacaacctgtggagaaactcctagaccacacgtta   | 4320 |
| OY | 4591 | ctgactgcctacacttcogaatccaaccccagaagcttggtgacccaagaattgccccttaagatg | 4650 |
| Db | 4321 | ctgactgcctacacttcogaatccaaccccagaagcttggtgacccaagaattgccccttaagatg | 4380 |
| OY | 4651 | gaaggtctctggctctggagacacagaactctactgactcgagcgaatctc 4699           |      |
| Db | 4381 | gaaggtctctggctctggagacacagaactctactgactcgagcgaatctc 4429           |      |

|          |                                  |
|----------|----------------------------------|
| RESULT   | 4                                |
| AAV23339 |                                  |
| ID       | AAV23339 standard; DNA; 4670 BP. |
| XX       |                                  |
| AC       | AAV23339;                        |
| XX       |                                  |
| DT       | 17-AUG-1998 (first entry)        |
| XX       |                                  |

|                            |                       |  |               |  |  |
|----------------------------|-----------------------|--|---------------|--|--|
| DE                         |                       |  |               |  | Human Factor-VIII gene lacking central B domain. |
| XX                         | KW                    | Factor-VIII; blood clotting; human; synthetic gene; codon usage; ss.   |               |  |  |
| XX                         | OS                    | Homo sapiens.  |               |  |  |
| XX                         | OS                    | Synthetic.   |               |  |  |
| XX                         | PN                    | M09812207-A1.  |               |  |  |
| XX                         | PD                    | 26-MAR-1998.   |               |  |  |
| XX                         | PF                    | 18-SEP-1997;   | 97WO-US16639. |  |  |
| XX                         | PR                    | 20-SEP-1996;   | 96US-071294.  |  |  |
| XX                         | PA                    | (GEHO ) GEN HOSPITAL CORP.   |               |  |  |
| XX                         | P1                    | Haas J, Seed B:  |               |  |  |
| XX                         | DR                    | WPI, 1998-21/200/19.   |               |  |  |
| XX                         | PT                    | New synthetic eukaryotic gene(s) - in which non-preferred or less preferred codon(s) are replaced to provide high level expression in mammalian cell(s)  |               |  |  |
| PS                         | XX                    | Example 3; Fig 12; 92pp: English.  |               |  |  |
| CC                         | CC                    | This gene codes for a human Factor-VIII protein that lacks the central B domain (amino acids 760-1639) of the native protein. In CC<br>a novel, claimed synthetic gene (see AAV23288), non-preferred or CC<br>less preferred codons of the native gene are replaced by codons CC<br>favored by highly expressed human genes to provide high-level CC<br>expression in mammalian cells. The synthetic gene was assembled from 29 pairs of oligonucleotides (see AAV23340-97) which served as PCR templates. Synthetic genes of the invention (see also CC<br>AAV23289-91) are used for production of recombinant proteins in CC<br>mammalian cells at levels of at least 500% of those obtained using the natural genes. They can also be used in gene therapy. |               |  |  |
| CC                         | CX                    | Sequence 4670 BP; 1325 A; 1072 C; 1036 G; 1237 T; 0 other;   |               |  |  |
| SQL                        |                       |  |               |  |  |
| Query Match                | Best Local Similarity | 91.3%; Score 4351.2; DB 19; Length 4670;   |               |  |  |
| Matches 4359; Conservative | 99.7%; Pred. No. 0;   | Mismatches 13; Indels 0; Gaps 0;   |               |  |  |
| OY                         | 328                   | gccaccagaagatactactcgtgtagtggaacatgcatacggtatatcaaaagtcat  | 387           |  |  |
| DB                         | 84                    | gccaccagaagatactactccgtggycagygaaactgltcatayggattataycaaagyat  | 143           |  |  |
| OY                         | 388                   | ctcgtgtgagtcgtctgttgnacgcaaatlctctctaagatgaccaaaacltttccatic   | 447           |  |  |
| DB                         | 144                   | ctcgtgtgagtcgtccgtgtgacgcaaatctctctctaagatgaccaaaacltttccatic  | 203           |  |  |
| OY                         | 448                   | aacctaccatgcctgtgttaaaaaagaactcttgtttaagaattcaagggtcaccttttacaac   | 507           |  |  |
| DB                         | 204                   | aacctaccatgcctgtgttaaaaaagaactcttgtttaagaattcaagggtcaccttttacaac   | 263           |  |  |
| OY                         | 508                   | atcgtctaagccaagcccacccctgtgatggctgtgtcgaagtccttacatccaagcggagtt  | 567           |  |  |
| DB                         | 264                   | atcgtctaagccaagcccacccctgtgatggctgtgtcgaagtccttacatccaagcggagtt  | 323           |  |  |
| OY                         | 568                   | tatgatacagtggttcattatcaactaaagacatygtcttccactctgttcagtcttcattgct,  | 627           |  |  |
| DB                         | 324                   | tatgatacagtggttcattatcaactaaagacatygtcttccactctgttcagtcttcattgct   | 383           |  |  |
| OY                         | 628                   | gttggttatcatctctcgtgaaagctcttcbaaggaagctgaataatgattgattagacaagatcaa  | 687           |  |  |
| DB                         | 384                   | gttggttatcatctcctaacggaagctcttcbaaggaagctgaataatgattgattagacaagatcaa   | 443           |  |  |
| OY                         | 688                   | aggagagaagaagaatgtataaagcttccctctgtgtgaagccatatatgtctggcagctc  | 747           |  |  |
| DB                         | 444                   | aggagagaagaagaatgtataaagcttccctctgtgtgaagccatatatgtctggcagctc  | 503           |  |  |

QY 748 ctgaagaagatggtcccaatgacctgaccacatgtgaccttaactcaatcttct 807  
|||||  
Db 504 ctgaagaagatggtcccaatgacctgaccacatgtgaccttaactcaatcttct 563  
QY 808 catgtgacctgtgttaaaagactgaatcagagccatctggagccctactatgtatga 867  
|||||  
Db 564 catgtgacctgtgttaaaagactgaatcagagccctactatgtatgtatga 623  
QY 868 gaaggagatctggccaaggaagaagacagaccccttgcaaaaatttaactacttttgc 927  
|||||  
Db 624 gaaggagatctggccaaggaagaagacagaccccttgcaaaaatttaactacttttgc 683  
QY 928 gtaattgataagaggaaagtctggcacctcagaaacaaagactccttgatgcagatag 987  
|||||  
Db 684 gtaattgataagaggaaagtctggcacctcagaaacaaagactccttgatgcagatag 743  
QY 988 gatgctgcacatctgtcgggacctggcctaaatgacacagtcgaatgttatgtataacagg 1047  
|||||  
Db 744 gatgctgcacatctgtcgggacctggcctaaatgacacagtcgaatgttatgtataacagg 803  
QY 1048 tctctccaaagtctgatatgtgatgacaggaatcagttctatggcatgtgattggaatg 1107  
|||||  
Db 804 tctctccaaagtctgatatgtgatgacaggaatcagttctatggcatgtgattggaatg 863  
QY 1108 ggcacacactcctgaagtgcacatcaatactcctcgaaggctcacacatctctgtgaggaa 1167  
|||||  
Db 864 ggcacacactcctgaagtgcacatcaatactcctcgaaggctcacacatctctgtgaggaa 923  
QY 1168 catcgccagagcgtctcctgtgaaactctgcgaataacttctctactatgctcaaacctctg 1227  
|||||  
Db 924 catcgccagagcgtctcctgtgaaactctgcgaataacttctctactatgctcaaacctctg 983  
QY 1228 atggaaccttggaacagttctcactgattgttgcatatctctctcccaacacatgagacatg 1287  
|||||  
Db 984 atggaaccttggaacagttctcactgattgttgcatatctctctcccaacacatgagacatg 1043  
QY 1288 gaagcttatgtcaaaagttagacagctgtccagaggaaaccccaactacgaatgaaaaataat 1347  
|||||  
Db 1044 gaagcttatgtcaaaagttagacagctgtccagaggaaaccccaactacgaatgaaaaataat 1103  
QY 1348 gaagaagcggaaagactatgatatgatacttactatctctgaatgtgattgttcaggtttc 1407  
|||||  
Db 1104 gaagaagcggaaagactatgatatgatacttactatctctgaatgtgattgttcaggtttc 1163  
QY 1408 gatgatagcacactcctcctcttatccaaatcgctcagttgccaagaaacatccctaa 1467  
|||||  
Db 1164 gatgatagcacactcctcctcttatccaaatcgctcagttgccaagaaacatccctaa 1223  
QY 1468 acttgggtacattacatctgtcgtcgaaggaggagactgggactatgctcccttagctctc 1527  
|||||  
Db 1224 acttgggtacattacatctgtcgtcgaaggaggagactgggactatgctcccttagctctc 1283  
QY 1528 gcccccgatgacagaggttataaagtcaaatattgaaacaaatggccctcagcggattggt 1587  
|||||  
Db 1284 gcccccgatgacagaggttataaagtcaaatattgaaacaaatggccctcagcggattggt 1343  
QY 1588 aggaagtacaaaaaagctccgatttatgtgcatacacagatgaaaccttaagaacgctgtaa 1647  
|||||  
Db 1344 aggaagtacaaaaaagctccgatttatgtgcatacacagatgaaaccttaagaacgctgtaa 1403  
QY 1648 gctattcagcatgaatcagaatcttgggaccttacttctatggggaagtgtgagaacaa 1707  
|||||  
Db 1404 gctattcagcatgaatcagaatcttgggaccttacttctatggggaagtgtgagaacaa 1463  
QY 1708 cgtgttatataatttaagaatccaagcagacacataatacatctaccctcaaggatc 1767  
|||||  
Db 1464 cgtgttatataatttaagaatccaagcagacacataatacatctaccctcaaggatc 1523  
QY 1768 actgatagtccgtctctgtatctcaagagattacaaaaggtgtaaacatttgaagat 1827  
|||||  
Db 1524 actgatagtccgtctctgtatctcaagagattacaaaaggtgtaaacatttgaagat 1583

QY 1828 ttccaattctgcagagagaataatccaatataaatggaacatgacgtatagaagatgg 1887  
|||||  
Db 1584 ttccaattctgcagagagaataatccaatataaatggaacatgacgtatagaagatgg 1643  
QY 1888 ccaactaatatagatctctcgtgtgcttgacccgctattactctagtgttcgttaatatggag 1947  
|||||  
Db 1644 ccaactaatatagatctctcgtgtgcttgacccgctattactctagtgttcgttaatatggag 1703  
QY 1948 agagatcagcttcaaggaactatgacctctctcactcgtccacaaagaatctgtatag 2007  
|||||  
Db 1704 agagatcagcttcaaggaactatgacctctctcactcgtccacaaagaatctgtatag 1763  
QY 2008 caaagaggaaacccaataatgttcagacaaaggaaatgtcatcctggtttctgtatttgat 2067  
|||||  
Db 1764 caaagaggaaacccaataatgttcagacaaaggaaatgtcatcctggtttctgtatttgat 1823  
QY 2068 gagaacccaagctgtgtacccctacagaaatataaagacttctcccaatccaagtgtga 2127  
|||||  
Db 1824 gagaacccaagctgtgtacccctacagaaatataaagacttctcccaatccaagtgtga 1883  
QY 2128 gtgcagcttgagatccaagattccaagctccaacatacatgcaacagatccaatggctat 2187  
|||||  
Db 1884 gtgcagcttgagatccaagattccaagctccaacatacatgcaacagatccaatggctat 1943  
QY 2188 gttttgatagtttgacagttgtcaggtttgttgcatagagtgacatctca 2247  
|||||  
Db 1944 gttttgatagtttgacagttgtcaggtttgttgcatagagtgacatctca 2003  
QY 2248 agcatitggagacagactgacttctctctgtctctctctctgatatcacttcaaac 2307  
|||||  
Db 2004 agcatitggagacagactgacttctctctgtctctctctctgatatcacttcaaac 2063  
QY 2308 aaaaatgtctatgaaagacacacactcaacctatcccatctcagaagaaactgtctcatg 2367  
|||||  
Db 2064 aaaaatgtctatgaaagacacacactcaacctatcccatctcagaagaaactgtctcatg 2123  
QY 2368 tcgattgaaanaacccaagttctatgtatcttggtgtgcacaaactcgaactctggaacaa 2427  
|||||  
Db 2124 tcgattgaaanaacccaagttctatgtatcttggtgtgcacaaactcgaactctggaacaa 2183  
QY 2428 ggcatacgcgccttaccagaggttctcagttgtgacaagaacatggtgttatccagg 2487  
|||||  
Db 2184 ggcatacgcgccttaccagaggttctcagttgtgacaagaacatggtgttatccagg 2243  
QY 2488 gacagttatgaaagatattccaagatactgtcgaatgaataaacaatgccaattgaacaa 2547  
|||||  
Db 2244 gacagttatgaaagatattccaagatactgtcgaatgaataaacaatgccaattgaacaa 2303  
QY 2548 agcttctcccaagaattccaagaacaccttagcactagagcaaaaagcaatttaatgccaacca 2607  
|||||  
Db 2304 agcttctcccaagaattccaagaacaccttagcactagagcaaaaagcaatttaatgccaacca 2363  
QY 2608 ccagctctgaaacgcatacagcggaataactcgtactcctctcaagttagatccaagag 2667  
|||||  
Db 2364 ccagctctgaaacgcatacagcggaataactcgtactcctctcaagttagatccaagag 2423  
QY 2668 gaaattgactatgtatatacatatcagttgaaatgaaagaaagattttgacattat 2727  
|||||  
Db 2424 gaaattgactatgtatatacatatcagttgaaatgaaagaaagattttgacattat 2483  
QY 2728 gatagaatgaaataatccaagagcccccagcttccaagaagaacacagacatattatc 2787  
|||||  
Db 2484 gatagaatgaaataatccaagagcccccagcttccaagaagaacacagacatattatc 2543  
QY 2788 gctcagatggagagctcgtggaattatgggatagtgctcccaacatgttccaagaac 2847  
|||||  
Db 2544 gctcagatggagagctcgtggaattatgggatagtgctcccaacatgttccaagaac 2603  
QY 2848 agggctcagagttgacagtgctccctcagttccaagaagaattttccaagaaatttaagtat 2907  
|||||  
Db 2604 agggctcagagttgacagtgctccctcagttccaagaagaattttccaagaaatttaagtat 2663  
QY 2908 ggtccttactcagccctatacagctgtgaggaactaaatgaacattgggactcctcggg 2967

|||||  
Db 2664 ggtccttactcagcccttataccgtggaacataatgacatttggagctctcctggg 2723  
QY 2968 ccatataaagaagcaagattggaagataataatcatgttaacttcagaatacagctctc 3027  
Db 2724 ccatataaagaagcaagattggaagataataatcatgttaacttcagaatacagctctc 2783  
QY 3028 gctccacttctctctatcttagccttattcttataggaagatacagaagcaaggagca 3087  
Db 2784 gctccacttctctctatcttagccttattcttataggaagatacagaagcaaggagca 2843  
QY 3088 gaacctagaanaaaacttgttcaagcctaataagcaaaacttaacttcttggaaagtcaa 3147  
Db 2844 gaacctagaanaaaacttgttcaagcctaataagcaaaacttaacttcttggaaagtcaa 2903  
QY 3148 catcataatgacacccactaaagaatgatttgaactgcaaaagcctgggttattctctgat 3207  
Db 2904 catcataatgacacccactaaagaatgatttgaactgcaaaagcctgggttattctctgat 2963  
QY 3208 gttagacctggaanaagattgttcaactcagccttgaatgagccctctggtctgcaact 3267  
Db 2964 gttagacctggaanaagattgttcaactcagccttgaatgagccctctggtctgcaact 3023  
QY 3368 aacacactgaaacctgttcatatgaggacaagtgaacagtaacagaaattgtctgttttc 3327  
Db 3024 aacacactgaaacctgttcatatgaggacaagtgaacagtaacagaaattgtctgttttc 3083  
QY 3328 accactcttgaatgagaccaaagctgttactcaactgaataatgaaagaactgcagg 3387  
Db 3084 accactcttgaatgagaccaaagctgttactcaactgaataatgaaagaactgcagg 3143  
QY 3388 gctcccgcaatatccagatggaagatcccaactttaagaagattatcgcttccatgca 3447  
Db 3144 gctcccgcaatatccagatggaagatcccaactttaagaagattatcgcttccatgca 3203  
QY 3448 atcaatggtctacataatgatactacacttgccttagtaatgctcagatcaaaagatt 3507  
Db 3204 atcaatggtctacataatgatactacacttgccttagtaatgctcagatcaaaagatt 3263  
QY 3508 cgaatggtatctgctcacaagcatgaggcaagaatgaaacatccatctatcttaagtga 3567  
Db 3264 cgaatggtatctgctcacaagcatgaggcaagaatgaaacatccatctatcttaagtga 3323  
QY 3568 catgtgttcaactgtacgaaaaaagaagatataaaatggcaactgtaacatctatcca 3627  
Db 3324 catgtgttcaactgtacgaaaaaagaagatataaaatggcaactgtaacatctatcca 3383  
QY 3628 ggtgttttggacagatggaatgttaccatccaaagcttgaatttggcggttgaatgc 3687  
Db 3384 ggtgttttggacagatggaatgttaccatccaaagcttgaatttggcggttgaatgc 3443  
QY 3688 ctatgtggcagcatctacatgtctggaatgagcaacttctctgtgtacagcaataag 3747  
Db 3444 ctatgtggcagcatctacatgtctggaatgagcaacttctctgtgtacagcaataag 3503  
QY 3748 tgtcagagctcccggtgaatggtcttctggacacatltagagattttagagcttca 3807  
Db 3504 tgtcagagctcccggtgaatggtcttctggacacatltagagattttagagcttca 3563  
QY 3808 ggaacaataatgagacatgtggccccaagcttggccagacttcaatltccggatccatcaat 3867  
Db 3564 ggaacaataatgagacatgtggccccaagcttggccagacttcaatltccggatccatcaat 3623  
QY 3868 gctctggagcacaagagcccttcttcttgatcaaggttgatctgttggcacaatgatt 3927  
Db 3624 gctctggagcacaagagcccttcttcttgatcaaggttgatctgttggcacaatgatt 3683  
QY 3928 attcagggcatcaagacccaaggtgtcccgctcaagaagtcttccaaacctctacatctcag 3987  
Db 3684 attcagggcatcaagacccaaggtgtcccgctcaagaagtcttccaaacctctacatctcag 3743  
QY 3988 ttatcatcatgtatagttctgtatggaagaagaatgagacattatcgagaagaattccact 4047  
|||||

Db 3744 ttatcatcatgtatagttctgtatggaagaagtgagcaacttatactgaggaattccact 3803  
QY 4048 ggaaccttaatggtcttcttcttggcaatgttgatttcatctggtgataaaacacatatttc 4107  
Db 3804 ggaaccttaatggtcttcttcttggcaatgttgatttcatctggtgataaaacacatatttc 3863  
QY 4108 aaacctccaattatgtctcgaatacatccggttggaccacaactatatagaattcagc 4167  
Db 3864 aaacctccaattatgtctcgaatacatccggttggaccacaactatatagaattcagc 3923  
QY 4168 actcttcgaatgagatgtggtctgtgatttaaatagttgcaagcatgccaatttggaaatg 4227  
Db 3924 actcttcgaatgagatgtggtctgtgatttaaatagttgcaagcatgccaatttggaaatg 3983  
QY 4228 gagaagtaagaatataatagatgacagatattactgtcttcaatcttaaccataatgttt 4287  
Db 3984 gagaagtaagaatataatagatgacagatattactgtcttcaatcttaaccataatgttt 4043  
QY 4288 gccacctgtctctcttcaaaaagctcgacttcaactcccaaggagagtaatgcttgaga 4347  
Db 4044 gccacctgtctctcttcaaaaagctcgacttcaactcccaaggagagtaatgcttgaga 4103  
QY 4348 cctcaggttgaataatccaaaagatggctgcaagtggacttccagaagacaatgaaatgctc 4407  
Db 4104 cctcaggttgaataatccaaaagatggctgcaagtggacttccagaagacaatgaaatgctc 4163  
QY 4408 acagagatgaactactcagggagtaaaatctctgcttaccagatgtatgtgaagagttc 4467  
Db 4164 acagagatgaactactcagggagtaaaatctctgcttaccagatgtatgtgaagagttc 4223  
QY 4468 ctcatctccagcagtaagaatggtcaccatgagtcctcttcttccagaatgcaaaagta 4527  
Db 4224 ctcatctccagcagtaagaatggtcaccatgagtcctcttcttccagaatgcaaaagta 4283  
QY 4528 aaggtttccaagggaatccaagactccttcaacactgtgtgaaactcctctagaccacag 4587  
Db 4284 aaggtttccaagggaatccaagactccttcaacactgtgtgaaactcctctagaccacag 4343  
QY 4588 ttactgactcgctacacttcttgatccaccccaaggttgggtgacccaagattgacctgag 4647  
Db 4344 ttactgactcgctacacttcttgatccaccccaaggttgggtgacccaagattgacctgag 4403  
QY 4648 atggaattcttgggtctgcgagacaggaaccttcaactgagctcgagcgagttc 4699  
Db 4404 atggaattcttgggtctgcgagacaggaaccttcaactgaggtgtgccaactgc 4455  
  
RESULT 5  
AAD00122  
ID AAD00122 standard; DNA: 4999 BP.  
XX  
AC AAD00122;  
XX  
DT 31-JUL-2000 (first entry)  
XX  
DE Recombinant adeno associated vector construct. pVmA.1cF8deltaB.  
XX  
KW Recombinant Adeno Associated Vector: rAAV; pVmA.1cF8deltaB; promoter;  
KW human factor VIII; hFVIII; EF1alpha; human elongation factor-1alpha;  
KW human growth hormone; hGH; ITR; inverted terminal repeat; haemophilias;  
KW gene therapy; ds.  
XX  
OS Adeno associated virus.  
XX  
PN WO200023116-A1.  
XX  
PD 27-APR-2000.  
XX  
PF 19-OCT-1999; 99NO-US24495.  
XX  
PR 20-OCT-1998; 98US-0104994.  
PR 24-MAR-1999; 99US-0125974.  
PR 30-JUL-1999; 99US-0364862.



[illegible]

|    |      |   |      |
|----|------|---|------|
| Dh | 891  | tcctbgaagaaagatggtcccaattggtccctctgacccactggtcccttaacttaactatctctt    | 950  |
| Qy | 806  | ctcatgtggaactctgtgtaaaaagacttgaattcaaggtctcatctggagacctactagtatgta    | 865  |
| Dh | 951  | ctctagctggaactctgtgtaaaaagacttgaattcaaggtccctcatctggagacctactagtatgta | 1010 |
| Qy | 866  | ggagaaagggaagctctgtgccaagaaagaaagacagaagccttgtaacaaattattractacttttg  | 925  |
| Dh | 1011 | ggagaaagggaagctctgtgccaagaaagaaagacagaagccttgtaacaaattattractacttttg  | 1070 |
| Qy | 926  | ctgtatcttgatgtaaggggaaaaaagtctggaactcagaacaaagaactctctgtatgtaacgata   | 985  |
| Dh | 1071 | ctgtatcttgatgtaaggggaaaaaagtctggaactcagaacaaagaactctctgtatgtaacgata   | 1130 |
| Qy | 986  | gggatgctgcatctgtctcgggtcctggtcccttaaatgcaacagtcnaatggttatgttaaca      | 1045 |
| Dh | 1131 | gggatgctgcatctgtctcgggtcctggtcccttaaatgcaacagtcnaatggttatgttaaca      | 1190 |
| Qy | 1046 | ggtctctcccaaggtcctgatctgtagtgcacaggaacaaactcatctggagctgtaattgaa       | 1105 |
| Dh | 1191 | ggtctctcccaaggtcctgatctgtagtgcacaggaacaaactcatctggagctgtaattgaa       | 1250 |
| Qy | 1106 | tgggacacacactctgtaaagtgcactccaatactctctgaaagtgccaacatctctgtbgaag      | 1165 |
| Dh | 1251 | tgggacacacacactctgtaaagtgcactccaatactctctgaaagtgccaacatctctgtbgaag    | 1310 |
| Qy | 1166 | acctatcgccaaaggtctcttggaatctcgcgcaataactcttctctactgtctaaacact         | 1225 |
| Dh | 1311 | acctatcgccaaaggtctcttggaatctcgcgcaataactcttctctactgtctaaacact         | 1370 |
| Qy | 1226 | tgaatgacctctggagcagtttctctaactgtttgtgtaatactctctcccaacacactgtgtgca    | 1285 |
| Dh | 1371 | tgaatgacctctggagcagtttctctaactgtttgtgtaatactctctcccaacacactgtgtgca    | 1430 |
| Qy | 1286 | tgggaagctcatgttcaaaagttagacagctgtctccagaaagaaaccccaactacgaatgtaaaaa   | 1345 |
| Dh | 1431 | tgggaagctcatgttcaaaagttagacagctgtctccagaaagaaaccccaactacgaatgtaaaaa   | 1490 |
| Qy | 1346 | atgaaagaagcggaaagactatgatatgatactctatctatctgtaaatgtgagtgtgagct        | 1405 |
| Dh | 1491 | atgaaagaagcggaaagactatgatatgatactctatctatctgtaaatgtgagtgtgagct        | 1550 |
| Qy | 1406 | ttgatgtagacaactctctctctctcttaatactcgtcagttgtgccaaagaaactccta          | 1465 |
| Dh | 1551 | ttgatgtagacaactctctctctctcttaatactcgtcagttgtgccaaagaaactccta          | 1610 |
| Qy | 1466 | aaactgtggtacatctatctatctgtcgtcgtgaaagaagaaagacgtggagacatctccctatgccc  | 1525 |
| Dh | 1611 | aaactgtggtacatctatctatctgtcgtcgtgaaagaagaaagacgtggagacatctccctatgccc  | 1670 |
| Qy | 1526 | tcgcccccgatgacagaagatataaaaaagtccaatatttgtaacaaatggtccctcagcggatg     | 1585 |
| Dh | 1671 | tcgcccccgatgacagaagatataaaaaagtccaatatttgtaacaaatggtccctcagcggatg     | 1730 |
| Qy | 1586 | gttaggaagtataaaaaagctccgattatgtagcatatcagaatgaataacatttaagactgtg      | 1645 |
| Dh | 1731 | gttaggaagtataaaaaagctccgattatgtagcatatcagaatgaataacatttaagactgtg      | 1790 |
| Qy | 1646 | aagctatctacacatgaaatcaggaactcttgtagacatttaatttatgtgggaagtgtgtagaca    | 1705 |
| Dh | 1791 | aagctatctacacatgaaatcaggaactcttgtagacatttaatttatgtgggaagtgtgtagaca    | 1850 |
| Qy | 1706 | caactgttgatataatttaagaatccaagcagacagacacataaacaactaactcaccacggaa      | 1765 |
| Dh | 1851 | caactgttgatataatttaagaatccaagcagacagacacataaacaactaactcaccacggaa      | 1910 |
| Qy | 1766 | tcactgtggtccggtccttttatccaaggaagattaccaaaaggtttaaacattttgaag          | 1825 |
| Dh | 1911 | tcactgtggtccggtccttttatccaaggaagattaccaaaaggtttaaacattttgaag          | 1970 |
| Qy | 1826 | atttccaattctgcccgggaaataatccaatataaatggtgcagtgtgtaagaagt              | 1885 |



Db 1971 atttccaattctgcagagaaatctcaataataatgacagtctgtagaagatg 2030  
QY 1886 ggcacaataatcagatccctcggtgctgacccgctattactcagtcttgtaatatg 1945  
Db 2031 ggcacaataatcagatccctcggtgctgacccgctattactcagtcttgtaatatg 2090  
QY 1946 agagagatctagcttcagactcattgacctctccatctcgtctaaagaatctgtag 2005  
Db 2091 agagagatctagcttcagactcattgacctctccatctcgtctaaagaatctgtag 2150  
QY 2006 atcaaaaggagaaacagataatgctcagacaagaaggaaatgcatcctgtcttctgtatttg 2065  
Db 2151 atcaaaaggagaaacagataatgctcagacaagaaggaaatgcatcctgtcttctgtatttg 2210  
QY 2066 atgagaacccgaagctgtgtactccaagagaaatatacagctttcttcccaatccagctg 2125  
Db 2211 atgagaacccgaagctgtgtactccaagagaaatatacagctttcttcccaatccagctg 2270  
QY 2126 gagtgcagcttgagaaatccagagcttccaaagcctccaacatcatgacagcatcaatgctc 2185  
Db 2271 gagtgcagcttgagaaatccagagcttccaaagcctccaacatcatgacagcatcaatgctc 2330  
QY 2186 atgtttttgtaagtttgcagttgtcagtttgttgcgaatgaggtgcgaactggtatcaatc 2245  
Db 2331 atgtttttgtaagtttgcagttgtcagtttgttgcgaatgaggtgcgaactggtatcaatc 2390  
QY 2246 taagcatctgagacaagacagactc 2305  
Db 2391 taagcatctgagacaagacagactctctctctctctctctctctctctctctctctctctc 2450  
QY 2306 acaaaatggtctatgaaagacacatccacatccatctccatctcagagaaactgtctc 2365  
Db 2451 acaaaatggtctatgaaagacacatccacatccatctccatctcagagaaactgtctc 2510  
QY 2366 tctcgaatgcaaaaacccaggtctatgattctcgtgggtgcacaactcagactcttcgagaca 2425  
Db 2511 tctcgaatgcaaaaacccaggtctatgattctcgtgggtgcacaactcagactcttcgagaca 2570  
QY 2426 gaagcatcagaccgctctacttaagaatttctagtgtgacaaagaacactggtatatacag 2485  
Db 2571 gaagcatcagaccgctctacttaagaatttctagtgtgacaaagaacactggtatatacag 2630  
QY 2486 aggaacagttaatgaatatctcagcatactgtcgtgaaataaacaatgcatctgaacaa 2545  
Db 2631 aggaacagttaatgaatatctcagcatactgtcgtgaaataaacaatgcatctgaacaa 2690  
QY 2546 gaagcttctccagaatctcaagaacccctagcacatgagcaaaagcaattaatgcaccc 2605  
Db 2691 gaagcttctccagaatctcaagaacccctagcacatgagcaaaagcaattaatgcaccc 2708  
QY 2606 caccagctctgaaagcgcatacgaagcaaaataactcgtactactcttcagtcagatcaag 2665  
Db 2709 caccagctctgaaagcgcatacgaagcaaaataactcgtactactcttcagtcagatcaag 2768  
QY 2666 aggaatctgactatgatacatatcagcttgaatgaagaaggaagatcttgcacatc 2725  
Db 2769 aggaatctgactatgatacatatcagcttgaatgaagaaggaagatcttgcacatc 2828  
QY 2726 atgataaggaatgaataatcagagcccccgcagcttccaaagaacaaacgacacathtta 2785  
Db 2829 atgataaggaatgaataatcagagcccccgcagcttccaaagaacaaacgacacathtta 2888  
QY 2786 ttctcgtcagatgaggaagcctctcgtgattatgagatgagtaactcccaacatgcttaagaa 2845  
Db 2889 ttctcgtcagatgaggaagcctctcgtgattatgagatgagtaactcccaacatgcttaagaa 2948  
QY 2846 acagggctcagagatgagcaggtccctcagttcaagaagaatgttcttccagaaatttaag 2905  
Db 2949 acagggctcagagatgagcaggtccctcagttcaagaagaatgttcttccagaaatttaag 3008  
QY 2906 atggtctccttaactcagccttatacgcgtgagagaactaaatgaacatcttggaactcctg 2965  
Db 3009 atggtctccttaactcagccttatacgcgtgagagaactaaatgaacatcttggaactcctg 3068

QY 2966 ggcacataataaagacgaaggttgaaagatatatactcagtgtaactttcagaataatcagcct 3025  
Db 3069 ggcacataataaagacgaaggttgaaagatatatactcagtgtaactttcagaataatcagcct 3128  
QY 3026 ctgcgtccctatctcccttactctcagccttattctctctgaggaagatcagaagcaagag 3085  
Db 3129 ctgcgtccctatctcccttactctcagccttattctctctgaggaagatcagaagcaagag 3188  
QY 3086 cagaacctaagaaaaaactctgtccaagcctaatagaaccaaacttactcttggaagatgc 3145  
Db 3189 cagaacctaagaaaaaactctgtccaagcctaatagaaccaaacttactcttggaagatgc 3248  
QY 3146 aacatacatggaaccccaaaagatgagtgttgcactgcaaaagcttggtctattctctg 3205  
Db 3249 aacatacatggaaccccaaaagatgagtgttgcactgcaaaagcttggtctattctctg 3308  
QY 3206 atgttgacctggaanaaagatgctcactcagagcctgtaattggaacccctctgtctgcaca 3365  
Db 3309 atgttgacctggaanaaagatgctcactcagagcctggtgatctggaacccctctgtctgcaca 3368  
QY 3266 ctaacacactggaacccctgctcctcagtgagagacaagtgcagtaacagaaattgctctgttc 3325  
Db 3369 ctaacacactggaacccctgctcctcagtgagagacaagtgcagtaacagaaattgctctgttc 3428  
QY 3326 tcaacacacttggatgagaccaaagcgtgtaacttcaactggaataatgagaaagaaactgca 3385  
Db 3429 tcaacacacttggatgagaccaaagcgtgtaacttcaactggaataatgagaaagaaactgca 3488  
QY 3386 gggctccctgcaatctccagatccagatccacacttcaagaagaattacacgtctccatg 3445  
Db 3489 gggctccctgcaatctccagatccagatccacacttcaagaagaattacacgtctccatg 3548  
QY 3446 caatcaatggtctacataatgataacactacccgtgctagtgaatggtcagaatccaagga 3505  
Db 3549 caatcaatggtctacataatgataacactacccgtgctagtgaatggtcagaatccaagga 3608  
QY 3506 ttctgagtgtatctgctcgaatgagtgagacgaatgaanaacacatcttcatctattctcag 3565  
Db 3609 ttctgagtgtatctgctcgaatgagtgagacgaatgaanaacacatcttcatctattctcag 3668  
QY 3566 gacatggtctcactgttcaacgaanaaagaagagatataaaatgagactgtacaactctatc 3625  
Db 3669 gacatggtctcactgttcaacgaanaaagaagagatataaaatgagactgtacaactctatc 3728  
QY 3626 caggtgttttttgagacagtgagaaatgttaccatccaagctvgaaatttggcgggtggaaat 3685  
Db 3729 caggtgttttttgagacagtgagaaatgttaccatccaagctvgaaatttggcgggtggaaat 3788  
QY 3686 gcttatttgccgaagatctcactcagcttgagtgagacacacttttctggtgtacaagcaata 3745  
Db 3789 gcttatttgccgaagatctcactcagcttgagtgagacacacttttctggtgtacaagcaata 3848  
QY 3746 agtgcagactccctctgagatgctctctgacacatctagaatattcaattacagctt 3805  
Db 3849 agtgcagactccctctgagatgctctctgacacatctagaatattcaattacagctt 3908  
QY 3806 caggacaataatggaacagtggtgcccacaaagctvgccagaaacttcatctccggaatcaatca 3865  
Db 3909 caggacaataatggaacagtggtgcccacaaagctvgccagaaacttcatctccggaatcaatca 3968  
QY 3866 atgccttgagacacaaaggaaccccttctctggaatcgaagtgagatctcgttgcccaatgga 3925  
Db 3969 atgccttgagacacaaaggaaccccttctctggaatcgaagtgagatctcgttgcccaatgga 4028  
QY 3926 ttattacaggtacaaagaccccaaggtgcccgttcagaagttctccagccttataatctctc 3985  
Db 4029 ttattacaggtacaaagaccccaaggtgcccgttcagaagttctccagccttataatctctc 4088  
QY 3986 agtttatcatcatgtatagttctgatagggaagaagtgcagaaacttatcgaagaaatcca 4045  
Db 4089 agtttatcatcatgtatagttctgatagggaagaagtgcagaaacttatcgaagaaatcca 4148

```
OY 4046 ctggaaacctaaatggtctctcttggcaatgtgatcatctcgggataaacaacatattc 4105
    |||
DB 4149 ctggaacctaatatgctctcttcttgcaatgtgatcatctcgggataaacaacatattc 4208
OY 4106 ttaacctcaattatgtcttgatcacatcgcttggaccacacatataatgatcga 4165
    |||
DB 4209 ttaacctcaattatgtcttgatcacatcgcttggaccacacatataatgatcga 4268
OY 4166 gcaactctgcatagtatgtagggctgtgatttaaatagttgcagcatgtcattggaa 4225
    |||
DB 4269 gcaactctgcatagtatgtagggctgtgatttaaatagttgcagcatgtcattggaa 4328
OY 4226 tggagagtaagaacatatacagatgcacagatgaactgtcttacttaccatattgt 4285
    |||
DB 4329 tggagagtaagaacatatacagatgcacagatgaactgtcttacttaccatattgt 4388
OY 4286 ttgcacctgtgtctcctcaaaagctgcactcctccaaaggagggatagtcctgga 4345
    |||
DB 4389 ttgcacctgtgtctcctcaaaagctgcactcctccaaaggagggatagtcctgga 4448
OY 4346 gacctcaggtgaataatccaaagaatgtgtcgaagtgaacttccagaagaacatgaag 4405
    |||
DB 4449 gacctcaggtgaataatccaaagaatgtgtcgaagtgaacttccagaagaacatgaag 4508
OY 4406 tcaacaggaagtaactactcagaaggataaaatctctgtctaccagcatgtatgtgaagagt 4465
    |||
DB 4509 tcaacaggaagtaactactcagaaggataaaatctctgtctaccagcatgtatgtgaagagt 4568
OY 4466 tccctcactcagcagatcgaatggccatctcagttgactcttttttagaatgtgcaag 4525
    |||
DB 4569 tccctcactcagcagatcgaatggccatctcagttgactcttttttagaatgtgcaag 4628
OY 4526 taaaggttttcagagggaatcaagactcctcacaacctgtgtgaactcttagaccac 4585
    |||
DB 4629 taaaggttttcagagggaatcaagactcctcacaacctgtgtgaactcttagaccac 4688
OY 4586 cgttactcgaactgcctacactcgaattcaacccccagagttgtgtgacagagattgcctga 4645
    |||
DB 4689 cgttactcgaactgcctacactcgaattcaacccccagagttgtgtgacagagattgcctga 4748
OY 4646 ggaatgaggtttctgtgctgcagagacagagacctactgactgagc 4693
    |||
DB 4749 ggaatgaggtttctgtgctgcagagacagagacctactgactgagc 4796

RESULT 6
AAx82259
ID AAX82259 standard; cDNA; 9164 BP.
XX
AC AAX82259;
XX
DT 18-AUG-1999 (first entry)
XX
DE Beta-domain deleted Factor VIII protein encoding gene (uncorrected).
XX
KW Factor VIII protein; gene modification; gene therapy; clinical disorder;
KW splicing pattern; RNA processing; gene regulation; beta-domain; human;
KW ss.
XX
OS Homo sapiens.
XX
PN WO929848-A1.
XX
PD 17-JUN-1999.
XX
PF 25-NOV-1998; 98MO-US25354.
XX
PR 16-JAN-1998; 98US-0071596.
XX
PR 05-DEC-1997; 97US-0067614.
XX
PA (IMMU-) IMMUNE RESPONSE CORP.
XX
PI Bidlingmaier S, Gonzales JEN, III CR, Yang CO;
```

```
XX WP1: 1999-385602/32.
DR P-PSDB; AAY21675.
XX
PT Genes and vectors exhibiting increased expression and novel splicing
PT patterns, useful for expression of, e.g. beta-domain deleted factor
PT VIII
XX
PS Example 2; Page 79-89; 123bp; English.
XX
CC The invention describes novel genes and vectors exhibiting increased
CC expression and novel splicing patterns. It provides a gene encoding a
CC Factor VIII protein, that comprises one or more consensus or near
CC consensus splice sites which have been corrected to increase expression.
CC The method, DNA sequences and expression vectors can be used to increase
CC the expression of a gene, especially a Factor VIII gene. Genes containing
CC modified 5' and/or 3' untranslated regions have optimized expression
CC levels and tissue-specific expression. The methods are used for
CC identification and correction of consensus splice sites, addition of
CC introns, optimization of 5' and 3' untranslated regions and increase in
CC cytoplasmic RNA accumulation. Hence the DNAs are useful in gene therapy
CC to treat a clinical disorder, to study RNA processing and/or gene
CC regulation. The present sequence represents an uncorrected version of
CC beta-domain deleted factor VIII protein encoding gene sequence (construct
CC pC1-2). This was used to develop a new coding sequence for beta-domain
CC deleted factor VIII protein by correcting the consensus splice sites.
CC
SQ
Sequence 9164 BP; 2451 A; 2170 C; 2081 G; 2462 T; 0 other;

Query Match 90.8%; Score 4325.6; DB 20; Length 9164;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 4383; Conservative 0; Mismatches 9; Indels 42; Gaps 1;

OY 266 ccacatctcaaatgagatgctccacactgctctcttctgtgctttcgattctgcttta 325
    |||
DB 1001 ccacatctcaaatgagatgctccacactgctctcttctgtgctttcgattctgcttta 1060
OY 326 gtgcacacagaagaatactactcgtgtgcagtgaactgtatcgtggaatataatgaagt 385
    |||
DB 1061 gtgcacacagaagaatactactcgtgtgcagtgaactgtatcgtggaatataatgaagt 1120
OY 386 atctcgtgtgagctgctgtgtgagcgaagattcctcctagaatgagccaaatctttccat 445
    |||
DB 1121 atctcgtgtgagctgctgtgtgagcgaagattcctcctagaatgagccaaatctttccat 1180
OY 446 tcaaacactcagtcgtgtacaaaagaactggttgtagaattcaaggttccactttca 505
    |||
DB 1181 tcaaacactcagtcgtgtacaaaagaactggttgtagaattcaaggttccactttca 1240
OY 506 acatcgtctaagccaagccactgtgagtgtctgtctaggtcctaccatccaggtgag 565
    |||
DB 1241 acatcgtctaagccaagccactgtgagtgtctgtctaggtcctaccatccaggtgag 1300
OY 566 ttatgatacagtggttattacacttaagaacatggttccatccctgttaagttcatg 625
    |||
DB 1301 ttatgatacagtggttattacacttaagaacatggttccatccctgttaagttcatg 1360
OY 626 ctgtgtgtgtatcctcactggaagcttctgtgagagctgtaatatgatacgaaccagtc 685
    |||
DB 1361 ctgtgtgtgtatcctcactggaagcttctgtgagagctgtaatatgatacgaaccagtc 1420
OY 686 aaaggagaaagaatgataaagtcttccctgtgtggaagccatataatgtctgagag 745
    |||
DB 1421 aaaggagaaagaatgataaagtcttccctgtgtggaagccatataatgtctgagag 1480
OY 746 tctgtgaagaagaatggttccaatggtccttgaccacatgtgcttactactatattt 805
    |||
DB 1481 tctgtgaagaagaatggttccaatggtccttgaccacatgtgcttactactatattt 1540
OY 806 ctcatgtgacctgtgtaaaagacttgaaatcagggccctcatgtgaagccttactagtatga 865
    |||
DB 1541 ctcatgtgacctgtgtaaaagacttgaaatcagggccctcatgtgaagccttactagtatga 1600
```

QY 866 gaagaaggagctctgccaagaaacacacagacctctgcacaaattatactattcttg 925  
|||||  
Db 1601 gaagaaggagctctgccaagaaacacacagacctctgcacaaattatactattcttg 1660  
QY 926 ctgtattctgaaggaagaaagttgacctcagaaacaaagaactctgtatgcagata 985  
|||||  
Db 1661 ctgtattctgaaggaagaaagttgacctcagaaacaaagaactctgtatgcagata 1720  
QY 986 gggaatctgcatctgctcgggcctggtctaaatctgcacacagtcaatgtgtatgtaaca 1045  
|||||  
Db 1721 gggaatctgcatctgctcgggcctggtctaaatctgcacacagtcaatgtgtatgtaaca 1780  
QY 1046 ggtctctgcaggtctgtatgtagtgcacaggaatcagttctatcttgacgtgtatggaa 1105  
|||||  
Db 1781 ggtctctgcaggtctgtatgtagtgcacaggaatcagttctatcttgacgtgtatggaa 1840  
QY 1106 tgggacacactcctggaagtgcacataatctctcgaaggtgcacacattctctgttgaaga 1165  
|||||  
Db 1841 tgggacacactcctggaagtgcacataatctctcgaaggtgcacacattctctgttgaaga 1900  
QY 1166 accatgcagagcgctccttggaaatctgcacaaataacttcttactgtctaaacactct 1225  
|||||  
Db 1901 accatgcagagcgctccttggaaatctgcacaaataacttcttactgtctaaacactct 1960  
QY 1226 tgaatgaccttgacacagttcttactgtttgttcatactctctccacacaaatgtagtga 1285  
|||||  
Db 1961 tgaatgaccttgacacagttcttactgtttgttcatactctctccacacaaatgtagtga 2020  
QY 1286 tggaaactatctgcaagtagacagctgtccagagagaaccccaactacacaaatgtaaaata 1345  
|||||  
Db 2021 tggaaactatctgcaagtagacagctgtccagagagaaccccaactacacaaatgtaaaata 2080  
QY 1346 atgaaagaacggaagactatgtagtgcacttactgtgaatctgtagtgtgaagt 1405  
|||||  
Db 2081 atgaaagaacggaagactatgtagtgcacttactgtgaatctgtagtgtgaagt 2140  
QY 1406 tgaatgatacactctctctctcttataccaaattcgctcagttgccaagaacatcccta 1465  
|||||  
Db 2141 tgaatgatacactctctctctcttataccaaattcgctcagttgccaagaacatcccta 2200  
QY 1466 aaacttggttacatcatctgctgctgaagagagagactggtcactatgctcccttagtcc 1525  
|||||  
Db 2201 aaacttggttacatcatctgctgctgaagagagagactggtcactatgctcccttagtcc 2260  
QY 1526 tgcgcccgaatgacagagattataaaatgcaataattgaaacaatggtccctcagcgagtct 1585  
|||||  
Db 2261 tgcgcccgaatgacagagattataaaatgcaataattgaaacaatggtccctcagcgagtct 2320  
QY 1586 gtaggaagtagcaaaaagtcgattatgagcacaagatgaaacattgaagactcgtg 1645  
|||||  
Db 2321 gtaggaagtagcaaaaagtcgattatgagcacaagatgaaacattgaagactcgtg 2380  
QY 1646 aagctatcagcatgatacagaatcttgggaacttacttataggggaagttgtagaaca 1705  
|||||  
Db 2381 aagctatcagcatgatacagaatcttgggaacttacttataggggaagttgtagaaca 2440  
QY 1706 caactgttgaattataattaaagaatcaagcaagcagacataatacattccctcacgaa 1765  
|||||  
Db 2441 caactgttgaattataattaaagaatcaagcaagcagacataatacattccctcacgaa 2500  
QY 1766 tcaactatgctcgtcttgattatcaagagatctacaaaagttgtaaaaacatttgaag 1825  
|||||  
Db 2501 tcaactatgctcgtcttgattatcaagagatctacaaaagttgtaaaaacatttgaag 2560  
QY 1826 atttccaaattctgcagagagaataatcaaaataataatgacagtagtgaagatg 1885  
|||||  
Db 2561 atttccaaattctgcagagagaataatcaaaataataatgacagtagtgaagatg 2620  
QY 1886 ggccaactaaatcagaatctcctgggtgcctgacccgctattacttagttcgttaataatg 1945  
|||||  
Db 2621 ggccaactaaatcagaatctcctgggtgcctgacccgctattacttagttcgttaataatg 2680

QY 1946 agagagatctagcttcaggaactcatcttgccctccctcatctgtctacaagaatctgtag 2005  
|||||  
Db 2681 agagagatctagcttcaggaactcatcttgccctccctcatctgtctacaagaatctgtag 2740  
QY 2006 atcaaaagagaacccagataatgtgcagacaagaagaaatgcatccgttttctgtattg 2065  
|||||  
Db 2741 atcaaaagagaacccagataatgtgcagacaagaagaaatgcatccgttttctgtattg 2800  
QY 2066 atgagaacccgaagctgtgtactctcacagagataataaagcttctccccaatccagctg 2125  
|||||  
Db 2801 atgagaacccgaagctgtgtactctcacagagataataaagcttctccccaatccagctg 2860  
QY 2126 gagtgcagcttggaggtccaagagttccaagcctccaacatcatgcacagatcaatggct 2185  
|||||  
Db 2861 gagtgcagcttggaggtccaagagttccaagcctccaacatcatgcacagatcaatggct 2920  
QY 2186 atgttttgaatgttgcagttgtcagttgtttgtcagtaggtgtgcatactgttacaatc 2245  
|||||  
Db 2921 atgttttgaatgttgcagttgtcagttgtttgttgcagtaggtgtgcatactgttacaatc 2980  
QY 2246 taagcatgtgagacagagactgaacttcccttctgtctcttctctgtatatacttacaac 2305  
|||||  
Db 2981 taagcatgtgagacagagactgaacttcccttctgtctcttctctgtatatacttacaac 3040  
QY 2306 acaaaatgtctatgtagagacaacatcccatctccattccatgaggaagaactgtctca 2365  
|||||  
Db 3041 acaaaatgtctatgtagagacaacatcccatctccattccatgaggaagaactgtctca 3100  
QY 2366 tgtcgaatgaaaaacccaggtctatgtgattctgtgggtgcacaaatcagaacttgcgaaca 2425  
|||||  
Db 3101 tgtcgaatgaaaaacccaggtctatgtgattctgtgggtgcacaaatcagaacttgcgaaca 3160  
QY 2426 gaggcattgacgcgcttactcgaaggttcttagttgtgtacaagaacactgtgtatataag 2485  
|||||  
Db 3161 gaggcattgacgcgcttactcgaaggttcttagttgtgtacaagaacactgtgtatataag 3220  
QY 2486 aggaacgttatgtagaattatccagcatctgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2545  
|||||  
Db 3221 aggaacgttatgtagaattatccagcatctgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3280  
QY 2546 gaagctctcccaagaattcaagaacacccctagcacatgacaaagaacaaattaatgcaacc 2605  
|||||  
Db 3281 gaagctctcccaagaattcaagaacacacccctagcacatgacaaagaacaaattaatgcaacc 3298  
QY 2606 caccagcttctgaaagccatcaacggaagaataaactgttactatcttcagtcagatcaag 2665  
|||||  
Db 3299 caccagcttctgaaagccatcaacggaagaataaactgttactatcttcagtcagatcaag 3358  
QY 2666 aggaattgactatgtagtatacattacagttgaaatgaaaggaagaaattttgacattt 2725  
|||||  
Db 3359 aggaattgactatgtagtatacattacagttgaaatgaaaggaagaaattttgacattt 3418  
QY 2726 atgattgaggaatgaaaaatcagagcccccgaagcttcaaaaagaacacagacatattta 2785  
|||||  
Db 3419 atgattgaggaatgaaaaatcagagcccccgaagcttcaaaaagaacacagacatattta 3478  
QY 2786 tgcgtcagttgagagagctcctggattatgtagttagtgcctcccaatggttctaaga 2845  
|||||  
Db 3479 tgcgtcagttgagagagctcctggattatgtagttagtgcctcccaatggttctaaga 3538  
QY 2846 acagagctcagaggtgtagtgcctccctcagttcaagaagaatttttccaggaattacag 2905  
|||||  
Db 3539 acagagctcagaggtgtagtgcctccctcagttcaagaagaatttttccaggaattacag 3598  
QY 2906 atggtccttactcagcccttataacgctgtagaactaaatgtaacaatttggagctctgg 2965  
|||||  
Db 3599 atggtccttactcagcccttataacgctgtagaactaaatgtaacaatttggagctctgg 3658  
QY 2966 ggccatatataagagagaagttgtagaataatacatgttgaactttccaagaatcagcct 3025  
|||||  
Db 3659 ggccatatataagagagaagttgtagaataatacatgttgaactttccaagaatcagcct 3718  
QY 3026 ctgcctccattctctctatcttagccttatcttctatgaggaagatcagaggaagag 3085

|    |      |   |      |
|----|------|---|------|
| Db | 3719 | ctctgtccattctctctatctctacgcttatcttcttatctgaggaaatccagaggaaagag      | 3778 |
| Qy | 3086 | cagaacctagaaaaaaccttctgtcaagcctaa tgaaccaaaacttaccttctgaaag tgc     | 3145 |
| Db | 3779 | cagaacctagaaaaaaccttctgtcaagcctaa tgaaccaaaacttaccttctgaaag tgc     | 3838 |
| Qy | 3146 | aacatcatatgagaccacacaaagatgagtttga tctgcacaagcccttggtctattctctg     | 3205 |
| Db | 3839 | aacatcatatgagaccacacacaaagatgagtttga tctgcacaagcccttggtctattctctg   | 3898 |
| Qy | 3206 | atgttgacctggaaaaaagatgtgcactcaagcctgatcttgaaacctctgtctgcaca         | 3265 |
| Db | 3899 | atgttgacctggaaaaaagatgtgcactcaagcctgatcttgaaacctctgtctgcaca         | 3958 |
| Qy | 3266 | ctaaacctatgaaacctgtctcatctggggagaaagatgaaatcagacaggaatttgctctgttc   | 3325 |
| Db | 3959 | ctaaacctatgaaacctgtctcatctggggagaaagatgaaatcagacaggaatttgctctgttc   | 4018 |
| Qy | 3326 | tcaacca tcttttga tgaagaccaaaagctgtgcacttcatctgaaatatatgaaagaaactgca | 3385 |
| Db | 4019 | tcaacca tcttttga tgaagaccaaaagctgtgcacttcatctgaaatatatgaaagaaactgca | 4078 |
| Qy | 3386 | gggctccctgtcaatlatccaga tggaaagatcccaactttaagagaattatccgtccatg      | 3445 |
| Db | 4079 | gggctccctgtcaatlatccaga tggaaagatcccaactttaagagaattatccgtccatg      | 4138 |
| Qy | 3446 | catatcaatgtctatcaatgatgtatcaacttactgtttagaaatgtgctccagatcaagaag     | 3505 |
| Db | 4139 | catatcaatgtctatcaatgatgtatcaacttactgtttagaaatgtgctccagatcaagaag     | 4198 |
| Qy | 3506 | tctcgatgtatctctgtccagacatgtggcagcaatgaaaaatccatctctatctatctcatg     | 3565 |
| Db | 4199 | tctcgatgtatctctgtccagacatgtggcagcaatgaaaaatccatctctatctatctcatg     | 4258 |
| Qy | 3566 | ggcagatgtctcaactgtgtagaaaaaagagatgataaatgtgcaatgttaaatctctatc       | 3625 |
| Db | 4259 | ggcagatgtctcaactgtgtagaaaaaagagatgataaatgtgcaatgttaaatctctatc       | 4318 |
| Qy | 3626 | caggtgtcttcttgagacaagtggaaatcttacaatccaaaagcttggaatttggcgggtggaat   | 3685 |
| Db | 4319 | caggtgtcttcttgagacaagtggaaatcttacaatccaaaagcttggaatttggcgggtggaat   | 4378 |
| Qy | 3686 | ggccttatgtgcgagcatctacaatgctctgggagatgagcaacttttctgtgttacaagcata    | 3745 |
| Db | 4379 | ggccttatgtgcgagcatctacaatgctctgggagatgagcaacttttctgtgttacaagcata    | 4438 |
| Qy | 3746 | agttgcaagctcccccgtggagatgtctcttgaaacaatcttaagaatttccagatbaacat      | 3805 |
| Db | 4439 | agttgcaagctcccccgtggagatgtctcttgaaacaatcttaagaatttccagatbaacat      | 4498 |
| Qy | 3806 | caggacaatatgagacagtggccccaaaagctgtggccagacttcattatccggtacaatca      | 3865 |
| Db | 4499 | caggacaatatgagacagtggccccaaaagctgtggccagacttcattatccggtacaatca      | 4558 |
| Qy | 3866 | atggccctggagaccacaaggagcccttcttcttgatcgaatcgaagctgtgtgcaccaatga     | 3925 |
| Db | 4559 | atggccctggagaccacaaggagcccttcttcttgatcgaatcgaagctgtgtgtgcaccaatga   | 4618 |
| Qy | 3926 | ttatitcaagcatcaagagaccacaagggtgcccgtcaagaagtcttccagagccttaca tctctc | 3985 |
| Db | 4619 | ttatitcaagcatcaagagaccacaagggtgcccgtcaagaagtcttccagagccttaca tctctc | 4678 |
| Qy | 3986 | agttatcatcatatgatagtctctgtatgaggagaagatgtgcagacttatccaggaattcca     | 4045 |
| Db | 4679 | agttatcatcatatgatagtctctgtatgaggagaagatgtgcagacttatccaggaattcca     | 4738 |
| Qy | 4046 | cgggaaccttaatgtgctctctttgtggcaatgttgatattcatcttggaataaaacacaaattt   | 4105 |
| Db | 4739 | cgggaaccttaatgtgctctctttgtggcaatgttgatattcatcttggaataaaacacaaattt   | 4798 |
| Qy | 4106 | ttaaacctccaattatgtctgcatacatccgtttgcacccaactcatctatagaatttgc        | 4165 |

|          |   |  |      |
|----------|---|--|------|
| Db       | 4799  | ttaacccccaattatgtctgcgtacatacccggtttgcacccaactaatatattagcttcgca    | 4858 |
| Qy       | 4166  | gaactcttcgcatactgagtttgatgtggtcgtgatttaatatagttgcagacatgcatctgggaa | 4225 |
| Db       | 4859  | gaactcttcgcatactgagtttgatgtggtcgtgatttaatatagttgcagacatgcatctgggaa | 4918 |
| Qy       | 4226  | tggaagataaagacatatcgaatcacagattactgtctatcctcacttaccaatatgt         | 4285 |
| Db       | 4919  | tggaagataaagacatatcgaatcacagattactgtctatcctcacttaccaatatgt         | 4978 |
| Qy       | 4286  | ttgcacacttggtctccttcaaaagcttcgactcaacctccaaggagaglaatgccttga       | 4345 |
| Db       | 4979  | ttgcacacttggtctccttcaaaagcttcgactcaacctccaaggagaglaatgccttga       | 5038 |
| Qy       | 4346  | gaacctcaagtataatccaanaagatggtcgaagtgaactccagaagacataatgaag         | 4405 |
| Db       | 5039  | gaacctcaagtataatccaanaagatggtcgaagtgaactccagaagacataatgaag         | 5098 |
| Qy       | 4406  | tcacagagtgtaactactcagggataaactctctgtctaccagacatgtaagtgaagagt       | 4465 |
| Db       | 5099  | tcacagagtgtaactactcagggataaactctctgtctaccagacatgtaagtgaagagt       | 5158 |
| Qy       | 4466  | tcctcatctccagcaatcgaagaatggccatagtggactctctttttcagaatggcaag        | 4525 |
| Db       | 5159  | tcctcatctccagcaatcgaagaatggccatagtggactctctttttcagaatggcaag        | 5218 |
| Qy       | 4526  | taaaagtttttcagggagaatcaagaacctctcaacctgtgtggaactctcagaccac         | 4585 |
| Db       | 5219  | taaaagtttttcagggagaatcaagaacctctcaacctgtgtggaactctcagaccac         | 5278 |
| Qy       | 4586  | cgttactgaactgcgtactccttcgaattcaaccaccagaagtgggtgcaccagattgcctga    | 4645 |
| Db       | 5279  | cgttactgaactgcgtactccttcgaattcaaccaccagaagtgggtgcaccagattgcctga    | 5338 |
| Qy       | 4646  | ggatggaggttccttggtcgtcggagcagagacctctactgaactcggagcagatgc          | 4659 |
| Db       | 5339  | ggatggaggttccttggtcgtcggagcagagacctctactgaactcggagcagatgc          | 5392 |
| RESULT 7 |   |  |      |
| AAV19581 |   |  |      |
| ID       | AAV19581 standard; cDNA; 4832 BP.   |  |      |
| XX       | AAV19581;   |  |      |
| AC       |   |  |      |
| XX       |   |  |      |
| DT       | 06-AUG-1998 (first entry)   |  |      |
| XX       |   |  |      |
| DE       |   |  |      |
| XX       | Human factor VIII beta-domain deleted SQN deletion cDNA sequence.         |  |      |
| KM       | Replication defective: recombinant retrovirus; RRV, therapeutic protein;  |  |      |
| KM       | haemophilia; thrombosis; hypercoagulable disorder; liver disease; human;  |  |      |
| KM       | hepatitis; thalassemia; phenylketonuria; Lesch-Nyhan syndrome; diabetes;  |  |      |
| KM       | cystic fibrosis; Duchenne's Muscular Dystrophy; hypercholesterolemia;     |  |      |
| KM       | hypopituitarism; adenine deaminase deficiency; HIV infection; anaemia;    |  |      |
| KM       | Gaucher's syndrome; high blood pressure; Alzheimer's disease; autoimmune; |  |      |
| KW       | inflammatory disease; factor VIII; ss.                                    |  |      |
| XX       |   |  |      |
| OS       | Homo sapiens.   |  |      |
| XX       |   |  |      |
| PN       | W09800541-A2.   |  |      |
| XX       |   |  |      |
| PD       | 08-JAN-1998.  |  |      |
| XX       |   |  |      |
| XX       | 02-JUL-1997; 97WO-US11784.  |  |      |
| PF       |   |  |      |
| XX       |   |  |      |
| PR       | 04-JUN-1997; 97US-0869309.  |  |      |
| PR       | 03-JUL-1996; 96US-0645601.  |  |      |
| PR       | 13-AUG-1996; 96US-0696381.  |  |      |
| XX       |   |  |      |
| PA       | (CHIR ) CHIRON CORP.  |  |      |
| XX       |   |  |      |
| PI       | Allen JR, Barber JR, Boder M, Chang SMW, Chong K;                         |  |      |
| PI       | De la Vega D, Depoloni J, Greengard J, Hsu DC, Ibanez CE;                 |  |      |

PI Jolly DJ, Lee R, Mittelstaedt DM, Prussak CE, Respass JG,  
XX  
DR WPI: 1998-086966/08.  
DR P-PSDB: AAW46246.

PT New replication defective recombinant retro-viruses - which can be administered to provide long term systemic expression of therapeutic protein in blood, useful in, e.g. treating hyper-coagulable disorders

PS Example 28; Pages 210-213; 272pp; English.

CC - This CD encodes the beta-domain deleted SON deletion protein of human  
CC factor VIII. This is used in the construction of recombinant retroviral  
CC vectors expressing human factor VIII. The invention provides the  
CC preparation of replication defective recombinant retrovirus (RRV)  
CC expressing a therapeutic protein. The RRV preparation is resistant to  
CC degradation by human complement and is capable of inducing long term  
CC systemic expression of the therapeutic protein when administered  
CC intravenously to a human. The long term systemic expression results in a  
CC measurable level of the therapeutic protein being produced in the blood  
CC of the human for a period of at least 30 days after the administration of  
CC the RRV vector preparation. RRV's can be used for *in vivo* delivery of  
CC therapeutic protein to treat, e.g., haemophilia A, haemophilia B,  
CC thrombosis, hypercoagulable disorders, liver diseases such as hepatitis,  
CC disorders such as thalassemia, phenylketonuria, Lesch-Nyhan syndrome,  
CC severe combined immunodeficiency (SCID), cystic fibrosis, Duchenne's  
CC muscular dystrophy, inherited emphysema, familial hypercholesterolemia,  
CC diabetes, hypoparathyroidism, adenine deaminase deficiency, alpha-  
CC antitrypsin deficiency, Gaucher's syndrome, anaemia, infections such as  
CC HIV infection, high blood pressure, Alzheimer's disease, autoimmune or  
CC inflammatory disease or graft versus host disease. RRV's are capable of  
CC surviving inactivation in human serum thereby allowing efficient gene  
CC transfer over prolonged periods of time.

Sequence 4832 BP; 1381 A; 1089 C; 1055 G; 1307 T; 0 other;

|                            |       |              |       |                    |
|----------------------------|-------|--------------|-------|--------------------|
| Query Match                | 90.7% | Score 4321.6 | DB 19 | Length 4832        |
| Best Local Similarity      | 98.8% | Pred. No. 0  |       |                    |
| Matches 4379; Conservative | 0     | Mismatches   | 9     | Indels 42; Gaps 12 |

|    |     |   |     |
|----|-----|---|-----|
| QY | 270 | catgaataatgagatctccacccgctctcttcttgctgcttcttgagatctcgctttagtc       | 323 |
| Db | 71  | catgaataatgagatctccacccgctctcttcttgctgcttcttgagatctcgctttagtc       | 130 |
| QY | 330 | caccagaagaatactactccttggtgcagtcgagacgtctcaatggaactatagcaaatgatct    | 389 |
| Db | 131 | caccagaagaatactactccttggtgcagtcgagacgtctcaatggaactatagcaaatgatct    | 190 |
| QY | 390 | cggatgagctgctctgtgagcgcaagaattctcctctaagatgycacaatctttccattcaa      | 449 |
| Db | 191 | cggagagctgcctctgtgagcgcaagaattctcctctaagatgycacaatctttccattcaa      | 250 |
| QY | 450 | caactcgtcgtgtgacaagaagactcgtttgtagaatcaagttcaactttcaaat             | 509 |
| Db | 251 | caactcgtcgtgtgacaagaagactcgtttgtagaatcaagttcaactttcaaat             | 310 |
| QY | 510 | cgtctaagccaagcgccacccctctgagtcgtctcgtacagttcctacatccaggtctaggttta   | 569 |
| Db | 311 | cgtctaagccaagcgccacccctctgagtcgtctcgtacagttcctacatccaggtctaggttta   | 370 |
| QY | 570 | tgtatacagtggtctattactacttaagaacatgagcttcccatccgtcaagcttcatgctgt     | 629 |
| Db | 371 | tgtatacagtggtctattactacttaagaacatgagcttcccatccgtcaagcttcatgctgt     | 430 |
| QY | 630 | tgggtatctcactactggaagaactctctgagggagctctgaatgatgatgacacagccagttcaag | 689 |
| Db | 431 | tgggtatctcactactggaagaactctctgagggagctctgaatgatgatgacacagccagttcaag | 490 |
| QY | 690 | ggagaaagaagatgataaagtcttccctcgtgtggaagccatacatatgtctcgtgcagttcct    | 749 |
| Db | 491 | ggagaaagaagatgataaagtcttccctcgtgtggaagccatacatatgtctcgtgcagttcct    | 550 |

|    |      |   |      |
|----|------|---|------|
| OY | 750  | gaagagagatggtccaatgtgcctctgacccactgtgcttactactatcatcttccca          | 803  |
| Db | 551  | gaagagagatggtccaatgtgcctctgacccactgtgcttactactatcatcttccca          | 610  |
| OY | 810  | tgtgacacctgtgtaaaagacttgatctaaagccctcatcttggaagccctactgtatgtaga     | 869  |
| Db | 611  | tgtgacacctgtgtaaaagacttgatctaaagccctcatcttggaagccctactgtatgtaga     | 670  |
| OY | 870  | agggagctctgcccagaagaaaagacacagacctgtgacaaattatactactcttctgct        | 929  |
| Db | 671  | agggagctctgcccagaagaaaagacacagacctgtgacaaattatactactcttctgct        | 730  |
| OY | 930  | atttgatgaagaaggaaagtgttgacatccagaaaagaagaactccttgatgtcagaataga      | 989  |
| Db | 731  | atttgatgaagaaggaaagtgttgacatccagaaaagaagaactccttgatgtcagaataga      | 790  |
| OY | 990  | tgtctgcatctgtctcgtggcctcgtgacctaaatgacacagatcaatgttataagataagga     | 1049 |
| Db | 791  | tgtctgcatctgtctcgtggcctcgtgacctaaatgacacagatcaatgttataagataagga     | 850  |
| OY | 1050 | tctgcccaggtctgtgttgatgacgcccaggaatacagtctatgtgcatgtgtattggaaatgg    | 1109 |
| Db | 851  | tctgcccaggtctgtgttgatgacgcccaggaatacagtctatgtgcatgtgtattggaaatgg    | 910  |
| OY | 1110 | caccacctccctgaagatgacatcaatatctcccggaaggttcaacaattctgtgtgaagaaca    | 1169 |
| Db | 911  | caccacctccctgaagatgacatcaatatctcccggaaggttcaacaattctgtgtgaagaaca    | 970  |
| OY | 1170 | tccgccaagcgtcctctgtgaaaactccgcacaaatacttctctactgctcacaacactctgtat   | 1229 |
| Db | 971  | tccgccaagcgtcctctgtgaaaactccgcacaaatacttctctactgctcacaacactctgtat   | 1030 |
| OY | 1230 | ggacctgtggacagtttctactgttttggatcatctcttcccaacaatgatagatgata         | 1289 |
| Db | 1031 | ggacctgtggacagtttctactgttttggatcatctcttcccaacaatgatagatgata         | 1090 |
| OY | 1290 | agcttatgtcaaaatgaacacgctgtccagaaggaaaccccaactacgaatgtaaaaataatga    | 1349 |
| Db | 1091 | agcttatgtcaaaatgaacacgctgtccagaaggaaaccccaactacgaatgtaaaaataatga    | 1150 |
| OY | 1350 | agaaagcggaaagactatgatagtatccttactgattcttgaatgatagtgtgtccaaagtta     | 1409 |
| Db | 1151 | agaaagcggaaagactatgatagtatccttactgattcttgaatgatagtgtgtccaaagtta     | 1210 |
| OY | 1410 | tgaatgaacaactctccctcccttcttcaaaatctgactagttgcccagaagaacatccctaaac   | 1469 |
| Db | 1211 | tgaatgaacaactctccctcccttcttcaaaatctgactagttgcccagaagaacatccctaaac   | 1270 |
| OY | 1470 | tctggtatcaataactgtcgtcgtgaagaaggagactgtggaactatgtcccttagtccctgc     | 1529 |
| Db | 1271 | tctggtatcaataactgtcgtcgtgaagaaggagactgtggaactatgtcccttagtccctgc     | 1330 |
| OY | 1530 | ccccgcatgacagaaggttatataaagaatgaatttgaaacaatgtgcccctcagcggaattgtag  | 1589 |
| Db | 1331 | ccccgcatgacagaaggttatataaagaatgaatttgaaacaatgtgcccctcagcggaattgtag  | 1390 |
| OY | 1590 | gaagtatacaaaaagatcccgacttitaatgtgcatacacagatgaaaccttlaagactcgtgaagc | 1649 |
| Db | 1391 | gaagtatacaaaaagatcccgacttitaatgtgcatacacagatgaaaccttlaagactcgtgaagc | 1450 |
| OY | 1650 | tattcagcatgaaatcaggaatctctgggaaccttacttataatgggaagtgtgagacacact     | 1709 |
| Db | 1451 | tattcagcatgaaatcaggaatctctgggaaccttacttataatgggaagtgtgagacacact     | 1510 |
| OY | 1710 | gtgtatataatgaagaatcaagcagaacagacataataacatctccctcgaaggaataac        | 1769 |
| Db | 1511 | gtgtatataatgaagaatcaagcagaacagacataataacatctccctcgaaggaataac        | 1570 |
| OY | 1770 | tgaatccgtcctctgtatcatcaaggagattcccaaaaggtgtlaaacaatttgaagaatt       | 1829 |
| Db | 1571 | tgaatccgtcctctgtatcatcaaggagattcccaaaaggtgtlaaacaatttgaagaatt       | 1630 |

|    |      |  |       |
|----|------|--|-------|
| OY | 1830 | tcaatcttcgacggagaaatattcaaathtaaatgacagtgactgtagaagtgaggc        | 18839 |
| Db | 1631 | tcgaattccgcccggagaaatattccaataataaavgacagtgactgtagaagatggccc     | 16500 |
| OY | 1890 | aactaatcaagatccctcgtgctgcgcgaccgcatatactagatttcgttaatavgagag     | 19499 |
| Db | 1691 | aactaaatacgaatccctcgtgtgcttgaccgcgtatactcaagtttcgttaatavgagag    | 17500 |
| OY | 1950 | agatctagcttcagagatccatactgcccctccctcaatctgctacaaaagatctgtatga    | 20099 |
| Db | 1751 | agatctagcttcagagatccatactgcccctccctcaatctgctacaaaagatctgtatga    | 18100 |
| OY | 2010 | aagagggaaacacagataatgtcagacaaagaggaatgtcatcctgtttctgtatattgta    | 20559 |
| Db | 1811 | aagagggaaacacagataatgtcagacaaagaggaatgtcatcctgtttctgtatattgta    | 18700 |
| OY | 2070 | gaacggaagctgttactctcaagagaaatatacaagctttctcccaatccagcttgag       | 21259 |
| Db | 1871 | gaacggaagctgttactctcaagagaaatatacaagctttctcccaatccagcttgag       | 19300 |
| OY | 2130 | gaagctctgagatccagagatttccaagctcccaacatcatgacacagcatcaatgctatgt   | 21859 |
| Db | 1931 | gaagctctgagatccagagatttccaagctcccaacatcatgacacagcatcaatgctatgt   | 19900 |
| OY | 2190 | ttttgtagtttgcaagttgtcagttgtgttgatgagtgagcatcgtatcatlctaa         | 2249  |
| Db | 1991 | ttttgtagtttgcaagttgtcagttgtgttgatgagtgagcatcgtatcatlctaa         | 20500 |
| OY | 2250 | catctggaagacaaagacagactccctctctgtcctctctctcgtgatatccctcaacaa     | 23099 |
| Db | 2051 | catctggaagacaaagacagactccctctctgtcctctctctcgtgatatccctcaacaa     | 21100 |
| OY | 2310 | aatgtgctatagaagacacacatccacccctatcccatctcagaggaagaaactgtctcatgtc | 23659 |
| Db | 2111 | aatgtgctatagaagacacacatccacccctatcccatctcagaggaagaaactgtctcatgtc | 21700 |
| OY | 2370 | gatgtgaaaccccaaggtctatgtgattctgggtgtgccaaactcagaacttcggaaacag    | 24299 |
| Db | 2171 | gatgtgaaaccccaaggtctatgtgattctgggtgtgccaaactcagaacttcggaaacag    | 22300 |
| OY | 2430 | catgacggcctctcaagttctcagttgtgtgcacaaagaaactgtatatactacggga       | 24899 |
| Db | 2231 | catgacggcctctcaagttctcagttgtgtgcacaaagaaactgtatatactacggga       | 22900 |
| OY | 2490 | cagttatgaagatatctcaagcatactgtgcgagtaaaacaaatgccatctgacccaagaag   | 25499 |
| Db | 2291 | cagttatgaagatatctcagcatactgtgcgagtaaaacaaatgccatctgacccaagaag    | 23500 |
| OY | 2550 | ctctctccagaaatccaagacacccctagcactaagcnaaagcaalttaatgcccacc       | 26099 |
| Db | 2351 | ctctctccagaa-----cccaacc   | 23680 |
| OY | 2610 | atctctgaaagcgcatacaagggaaataactgttactaccccttcaagtcagatacgaagga   | 26659 |
| Db | 2369 | atctctgaaagcgcatacaagggaaataactgttactaccccttcaagtcagatacgaagga   | 24280 |
| OY | 2670 | aattgactatgatatccataccatcagtttgaatgaaagaaagaaatcttgacatttga      | 27299 |
| Db | 2429 | aattgactatgatatccataccatcagtttgaatgaaagaaagaaatcttgacatttga      | 24880 |
| OY | 2730 | tgaagatgaaaaatcagagcccccgcagcttccaagaagaaacacagacataatttatg      | 27899 |
| Db | 2489 | tgaagatgaaaaatcagagcccccgcagcttccaagaagaaacacagacataatttatg      | 25480 |
| OY | 2790 | tgcagttgaggaagctcttggtgataatggatgagtagctcccccacatgttctaagaacag   | 2849  |
| Db | 2549 | tgcagttgaggaagctcttggtgataatggatgagtagctcccccacatgttctaagaacag   | 26080 |
| OY | 2850 | ggctcaggaatggcagagtgccctcaagttccaagaagatgtgtttccaagaatttactgtg   | 29099 |
| Db | 2609 | ggctcaggaatggcagagtgccctcaagttccaagaagatgtgtttccaagaatttactgtg   | 26680 |
| OY | 2910 | ctccttactacagcccttataccgttgagaagacataaagaaatcttggaactccctgggccc  | 29699 |

|    |      |   |      |
|----|------|---|------|
| Db | 2669 | cccccttaacagcccttaatacgttggaagacaaagaaacatttggagctccctggggcc          | 2728 |
| Qy | 2970 | atataaagagcagaagtttgaagaataatacatggttaactttcaagaataacagagctctcg       | 3029 |
| Db | 2729 | atataaagagcagaagtttgaagaataatacatggttaactttcaagaataacagagctctcg       | 2788 |
| Qy | 3030 | tcctcatctcttatacttaagccttaatttcttatggaagaatcagaaggcaagagcaga          | 3089 |
| Db | 2789 | tcctcatctcttatacttaagccttaatttcttatggaagaatcagaaggcaagagcaga          | 2848 |
| Qy | 3090 | acctgaaaaaaacttttcaaagccttaatgaaaccaaacttaacttttgaagttgaca            | 3149 |
| Db | 2849 | acctgaaaaaaacttttcaaagccttaatgaaaccaaacttaacttttgaagttgaca            | 2908 |
| Qy | 3150 | lcaatatgccaccaactaaagaatgaattttgactgccaagccttggtgctattcttcgagtt       | 3209 |
| Db | 2909 | lcaatatgccaccaactaaagaatgaattttgactgccaagccttggtgctattcttcgagtt       | 2968 |
| Qy | 3210 | tgaaccttgaaaaaagatgctgacatcagccttgatctggaccctctggtctgcgacataa         | 3269 |
| Db | 2969 | tgaaccttgaaaaaagatgctgacatcagccttgatctggaccctctggtctgcgacataa         | 3028 |
| Qy | 3270 | caacaetgaaccttgctcaatggtggaacaaagtgaacatcagggaaatttgctctgttttaac      | 3329 |
| Db | 3029 | caacaetgaaccttgctcaatggtggaacaaagtgaacatcagggaaatttgctctgttttaac      | 3088 |
| Qy | 3330 | catctttgaatgaagcccaaaaagctgtgtaacttccacttgaataatataygaaagaacctgcaaggc | 3389 |
| Db | 3089 | catctttgaatgaagcccaaaaagctgtgtaacttccacttgaataatataygaaagaacctgcaaggc | 3148 |
| Qy | 3390 | tcacctgcaataatccaagatggaagaatcccaacttttaagaagaaatatactgccttccatgacaat | 3449 |
| Db | 3149 | tcacctgcaataatccaagatggaagaatcccaacttttaagaagaaatatactgccttccatgacaat | 3208 |
| Qy | 3450 | caatggcctaactaatgtgataaactacttggtcttagtaatggtgcagatccaagaattccg       | 3509 |
| Db | 3209 | caatggcctaactaatgtgataaactacttggtcttagtaatggtgcagatccaagaattccg       | 3268 |
| Qy | 3510 | atggaatactgttccagctatggtgcagccaatgaaacaatccattcattcattcaatgagga       | 3569 |
| Db | 3269 | atggaatactgttccagctatggtgcagccaatgaaacaatccattcattcattcaatgagga       | 3328 |
| Qy | 3570 | tgtgttccaactgtaagaaaaaaagagagataataaatgagcactgtataacatctatccacag      | 3629 |
| Db | 3329 | tgtgttccaactgtaagaaaaaaagagagataataaatgagcactgtataacatctatccacag      | 3388 |
| Qy | 3630 | tgtttttgagaacagttggaataatggtataccaatcgaagcttggaatttggcgggtggaatgctt   | 3689 |
| Db | 3389 | tgtttttgagaacagttggaataatggtataccaatcgaagcttggaatttggcgggtggaatgctt   | 3448 |
| Qy | 3690 | tatttgccaagatactataatccttggaatggaagacaacttttcttcggtgtataagaanaaatgtg  | 3749 |
| Db | 3449 | tatttgccaagatactataatccttggaatggaagacaacttttcttcggtgtataagaanaaatgtg  | 3508 |
| Qy | 3750 | tcagactccctctgggaatggtcttcttggaacaaatlaagaaatttcagatatacagcttcagg     | 3809 |
| Db | 3509 | tcagactccctctgggaatggtcttcttggaacaaatlaagaaatttcagatatacagcttcagg     | 3568 |
| Qy | 3810 | acaataatgagacagtggggtccccaagctgtgcagacttcatattccggatcaataatgc         | 3869 |
| Db | 3569 | acaataatgagacagtggggtccccaagctgtgcagacttcatattccggatcaataatgc         | 3628 |
| Qy | 3870 | cttgagacccaagaagacccctttcttcggaatcagaagtggaatctggttggaccaaaatattat    | 3929 |
| Db | 3629 | cttgagacccaagaagacccctttcttcggaatcagaagtggaatctggttggaccaaaatattat    | 3688 |
| Qy | 3930 | tcacggacataaagcccaagggtgcggctcagaagttcttcagccttaaatcttcagtt           | 3989 |
| Db | 3689 | tcacggacataaagcccaagggtgcggctcagaagttcttcagccttaaatcttcagtt           | 3748 |
| Qy | 3990 | tatcaatagtataagttcttgaatgggaagaatggtgacagattatctgaggaatctccatg        | 4049 |



[illegible]

|                            |  |               |
|----------------------------|--|---------------|
| PR                         | 04-JUN-1997;   | 97US-0869309. |
| PR                         | 03-JUL-1996;   | 96US-0645601. |
| PR                         | 13-AUG-1996;   | 96US-0696381. |
| XX                         |  |               |
| PA                         | (CHIR ) CHIRON CORP.   |               |
| XX                         |  |               |
| PI                         | Allen JR, Barber JR, Boder M, Chang SM, Chong K;                     |               |
| PI                         | De La Vega D, Depolo NJ, Greengard J, Hsu DC, Ibanez CE;             |               |
| PI                         | Jolly DJ, Mittlestaedt DM, Prussak CE, Respass JG;                   |               |
| XX                         |  |               |
| DR                         | WPI: 1998-086967/08.   |               |
| DR                         | P-PsDB: AAM44372.  |               |
| XX                         |  |               |
| PT                         | New replication defective recombinant retroviruses - which express B |               |
| PT                         | domain-deleted human factor VIII or human factor IX for the          |               |
| PT                         | treatment of haemophilia   |               |
| XX                         |  |               |
| PS                         | Claim 6; Page 174-175; 236pp; English.                               |               |
| XX                         |  |               |
| CC                         | This DNA sequence includes a coding region for the B domain          |               |
| CC                         | deletion mutant SON (see AAM44372) of human Factor VIII. The SON     |               |
| CC                         | mutant is created by fusing Ser-743 to Gln-1638 of native Factor     |               |
| CC                         | VIII (see AAM44373) to form a Ser-Gln-Asn (SON) link between the A2  |               |
| CC                         | and A3 Factor VIII domains. When compared to plasmid-derived         |               |
| CC                         | Factor VIII, the SON deletion does not influence the in vivo         |               |
| CC                         | pharmacokinetics, but the reduced size of the molecule appears to    |               |
| CC                         | decrease proteolytic degradation. The invention relates to           |               |
| CC                         | preparations of replication defective recombinant retrovirus (RV)    |               |
| CC                         | expressing a B domain-deleted human Factor VIII protein, where the   |               |
| CC                         | recombinant RV is capable of infecting human cells, is resistant     |               |
| CC                         | to degradation by human complement and is capable of inducing        |               |
| CC                         | long-term (at least 30 days and up to 6 months or longer             |               |
| CC                         | post-injection) systemic expression of Factor VIII when              |               |
| CC                         | administered to a haemophilia A patient.                             |               |
| XX                         |  |               |
| SQ                         | Sequence 4832 BP; 1381 A; 1089 C; 1055 G; 1307 T; 0 other;           |               |
| Query Match                | 90.7%; Score 4321.6; DB 19; Length 4832;                             |               |
| Best Local Similarity      | 98.8%; Pred. No. 0;  |               |
| Matches 4379; Conservative | 0; Mismatches 9; Indels 42; Gaps                                     |               |
| QY                         | 270 catgcaaatagagcctccaccctgctcttcttctgctgaccttgcattgcttaagtc        | 329           |
| DB                         | 71 catgcaaatagagcctccaccctgctcttcttctgctgaccttgcattgcttaagtc         | 130           |
| QY                         | 330 caccagaagatactactctggtgtgcagtgtaactgtcatatgagctatataagtgatct     | 389           |
| DB                         | 131 caccagaagatactactctggtgtgcagtgtaactgtcatatgagctatataagtgatct     | 190           |
| QY                         | 390 cgggtagctgctgtgtgacgcaagatcttcctctagatgtgccaaaatctttccattca      | 449           |
| DB                         | 191 cgggtagctgctgtgtgacgcaagatcttcctctagatgtgccaaaatctttccattca      | 250           |
| QY                         | 450 caccctcagtcgtgatacaaaaagacctgtgttttagaatcacaggttcacctttcaact     | 509           |
| DB                         | 251 caccctcagtcgtgatacaaaaagacctgtgttttagaatcacaggttcacctttcaact     | 310           |
| QY                         | 510 cgtctagccaaagccacacctgtgatactggtgtctgtagtccctaccatcacagtcgagttta | 569           |
| DB                         | 311 cgtctagccaaagccacacctgtgatactggtgtctgtagtccctaccatcacagtcgagttta | 370           |
| QY                         | 570 tgatacaagtgtcatctaatcaacttaagaacaatgtgtcccatcctctgcagtcttcaatgct | 629           |
| DB                         | 371 tgatacaagtgtcatctaatcaacttaagaacaatgtgtcccatcctctgcagtcttcaatgct | 430           |
| QY                         | 630 tgggtatctctactcgtggaagacctcttgagggagcgtbaatatgtagtgcagcaagccaag  | 689           |
| DB                         | 431 tgggtatctctactcgtggaagacctcttgagggagcgtbaatatgtagtgcagcaagccaag  | 490           |
| QY                         | 690 gtagagaagaagatgaataagcttcctcctgtgtggaagccatacatatgtctggcagtgctc  | 749           |
| DB                         | 491 gtagagaagaagatgaataagcttcctcctgtgtggaagccatacatatgtctggcagtgctc  | 550           |



|    |      |   |      |
|----|------|---|------|
| QY | 750  | gaaagagatggtccaaatgagctcttgacccactgtgcttaactactcaatattcttctca       | 809  |
| Db | 551  | gaaagagatggtccaaatgagctcttgacccactgtgcttaactactcaatattcttctca       | 610  |
| QY | 810  | tgtggaacctgtgaaagaacttgaaatctcaagcctcatgtggaacctactagatgtgagga      | 869  |
| Db | 611  | tgtggaacctgtgaaagaacttgaaatctcaagcctcatgtggaacctactagatgtgagga      | 670  |
| QY | 870  | aggagatcttgccaaggaagaacacagagacttgcacaatatctactacttttgcgt           | 929  |
| Db | 671  | aggagatcttgccaaggaagaacacagagacttgcacaatatctactacttttgcgt           | 730  |
| QY | 930  | atttgatgaagggaaaaagtgtgcaactcaagaacaagaacctccttgatgcagatagga        | 989  |
| Db | 731  | atttgatgaagggaaaaagtgtgcaactcaagaacaagaacctccttgatgcagatagga        | 790  |
| QY | 990  | tgtgtgcatctgctcgggccctgagccttaaatgacacacagatctgattgtgtaaaaggtc      | 1049 |
| Db | 791  | tgtgtgcatctgctcgggccctgagccttaaatgacacacagatctgattgtgtaaaaggtc      | 850  |
| QY | 1050 | tcttgcaagtgctgattggaatgagccaaagaaatcagctctatgtgcagtgtgaaatggg       | 1109 |
| Db | 851  | tcttgcaagtgctgattggaatgagccaaagaaatcagctctatgtgcagtgtgaaatggg       | 910  |
| QY | 1110 | caaccactccggaagtgtgacctccaatatctccttgaaaggtacacaaattctctgtgaggaacca | 1169 |
| Db | 911  | caaccactccggaagtgtgacctccaatatctccttgaaaggtacacaaattctctgtgaggaacca | 970  |
| QY | 1170 | tgcgcaagcgctctcttgaaatctcgcacaataactctccttaactgtgcaaacctcttat       | 1229 |
| Db | 971  | tgcgcaagcgctctcttgaaatctcgcacaataactctccttaactgtgcaaacctcttat       | 1030 |
| QY | 1230 | ggaaccttggaagcttctactggtttgtcatatctcttccacaacaatgagtgcagtga         | 1289 |
| Db | 1031 | ggaaccttggaagcttctactggtttgtcatatctcttccacaacaatgagtgcagtga         | 1090 |
| QY | 1290 | agattttgtcaaaagttaacagcgctgtccagaagaaaccccaactacgaatgaaataaatga     | 1349 |
| Db | 1091 | agattttgtcaaaagttaacagcgctgtccagaagaaaccccaactacgaatgaaataaatga     | 1150 |
| QY | 1350 | agaagcggaagactatgatagtatcttactgatacttgaaatgagatggtgcaggttga         | 1409 |
| Db | 1151 | agaagcggaagactatgatagtatcttactgatacttgaaatgagatggtgcaggttga         | 1210 |
| QY | 1410 | tgaatgacaactctcctcttattccaaattcgctcagttgccaagaagcatcctaaac          | 1469 |
| Db | 1211 | tgaatgacaactctcctcttattccaaattcgctcagttgccaagaagcatcctaaac          | 1270 |
| QY | 1470 | tctggtgacattatattgctgtgaagagggagacttggaacatgctcctctatgctctgc        | 1529 |
| Db | 1271 | tctggtgacattatattgctgtgaagagggagacttggaacatgctcctctatgctctgc        | 1330 |
| QY | 1530 | ccccgaatgaacagaagtataaaagtcaatatctgaaacaatgagcctcagcgagatgtgag      | 1589 |
| Db | 1331 | ccccgaatgaacagaagtataaaagtcaatatctgaaacaatgagcctcagcgagatgtgag      | 1390 |
| QY | 1590 | gaaagtacaaaaaaggtccgatttatgtgcatacacagatgaaacctttaagactcgtgaagc     | 1649 |
| Db | 1391 | gaaagtacaaaaaaggtccgatttatgtgcatacacagatgaaacctttaagactcgtgaagc     | 1450 |
| QY | 1650 | tattcagatgaaatcagaagatcttgagaccttiaacttaatgaggaaagtggagacaact       | 1709 |
| Db | 1451 | tattcagatgaaatcagaagatcttgagaccttiaacttaatgaggaaagtggagacaact       | 1510 |
| QY | 1710 | gttgattatataatgaatcaagaacagacagccaataacatactacccctcaggaatcac        | 1769 |
| Db | 1511 | gttgattatataatgaatcaagaacagacagccaataacatactacccctcaggaatcac        | 1570 |
| QY | 1770 | tgaatgctcgtctcttattatccaagggataatccaaaaaggtgtaaaaactttggaagattt     | 1829 |
| Db | 1571 | tgaatgctcgtctcttattatccaagggataatccaaaaaggtgtaaaaactttggaagattt     | 1630 |

|||||  
Db 2669 ctcccttaccagccctataccgttgagagactaaatgaacatttggactccctgggccc 2728  
QY 2970 atataataagcgagaagtgtgaagataataatcatatgttaacttcaagaatacagcctctcg 3029  
Db 2729 atataataagcgagaagtgtgaagataataatcatatgttaacttcaagaatacagcctctcg 2788  
QY 3030 tccctatctctctatctctagccttattctctatgaagaagatcagaagcgagaagcaga 3089  
Db 2789 tccctatctctctatctctagccttattctctatgaagaagatcagaagcgagaagcaga 2848  
QY 3090 acctagaaaaaacttctgcaagcctaaatgaacccaacttacttctggaagtgcacaa 3149  
Db 2849 acctagaaaaaacttctgcaagcctaaatgaacccaacttacttctggaagtgcacaa 2908  
QY 3150 tcatatgcccccaactaaagaatgatttgaactcgaaaagccttggtctattctctgagt 3209  
Db 2909 tcatatgcccccaactaaagaatgatttgaactcgaaaagccttggtctattctctgagt 2968  
QY 3210 tgacctgaaaaaagatgtgcactcaagcctgattggaacccctctctgtctgcacaaactaa 3269  
Db 2969 tgacctgaaaaaagatgtgcactcaagcctgattggaacccctctctgtctgcacaaactaa 3028  
QY 3270 cacactgaaccctgtctcatatggagacaaagtgaacagtaacaggaatttgcctctgtttcac 3329  
Db 3029 cacactgaaccctgtctcatatggagacaaagtgaacagtaacaggaatttgcctctgtttcac 3088  
QY 3330 catcttttgatgagacccaaagcctgttacttcaactgaaataatggaagaactgcagagc 3389  
Db 3089 catcttttgatgagacccaaagcctgttacttcaactgaaataatggaagaactgcagagc 3148  
QY 3390 tccctgcaatatccagaatggaagatcccaacttttaagaagaattatgccttccatgtaac 3449  
Db 3149 tccctgcaatatccagaatggaagatcccaacttttaagaagaattatgccttccatgtaac 3208  
QY 3450 caatggtacataatggaatacactacccttggtctagtaatgtcgcaggaatcaaaagattcg 3509  
Db 3209 caatggtacataatggaatacactacccttggtctagtaatgtcgcaggaatcaaaagattcg 3268  
QY 3510 atggtatctgtcagcatatggcgagcaatgaaacatccatctcatatcttcaagtgcaga 3569  
Db 3269 atggtatctgtcagcatatggcgagcaatgaaacatccatctcatatcttcaagtgcaga 3328  
QY 3570 tgtgtcacgtgacgaaaaaaagagagataaaatgycactgtacacatctctacacag 3629  
Db 3329 tgtgtcacgtgacgaaaaaaagagagataaaatgycactgtacacatctctacacag 3388  
QY 3630 tgtttttgagacagtggaatgttaccatccaaagctggaatttgcggttggaatgcct 3689  
Db 3389 tgtttttgagacagtggaatgttaccatccaaagctggaatttgcggttggaatgcct 3448  
QY 3690 tatttgcgagcatctacatagtcttgagatgagacaaacttttctgtgttaccagcaataagt 3749  
Db 3449 tatttgcgagcatctacatagtcttgagatgagacaaacttttctgtgttaccagcaataagt 3508  
QY 3750 tccagactcccccttggaatggtctctgacacatctagaagatttcagattacagcttcacag 3809  
Db 3509 tccagactcccccttggaatggtctctgacacatctagaagatttcagattacagcttcacag 3568  
QY 3810 acaatatgagcagtggtgccccaaagctgtgcagactcatattctcgatcaatcaatgc 3869  
Db 3569 acaatatgagcagtggtgccccaaagctgtgcagactcatattctcgatcaatcaatgc 3628  
QY 3870 ctgtgagcacaagaagccttcttctgtgacaaagtgtgactctgttgcacccaatgattat 3929  
Db 3629 ctgtgagcacaagaagccttcttctgtgacaaagtgtgactctgttgcacccaatgattat 3688  
QY 3930 tcaagcagcatcaagaccagagtgcccgtaagaagttctccagcctctacaatctcagtt 3989  
Db 3689 tcaagcagcatcaagaccagagtgcccgtaagaagttctccagcctctacaatctcagtt 3748  
QY 3990 tatcatcatgtatagttcttgaatggagaagatggcagacttatcgaagaattccactgg 4049  
|||||

Db 3749 tatcatcatgtatagttcttgaatggagaagatggcagacttatcgaagaattccactgg 3808  
QY 4050 aaccttaatgtcttcttcttggaatgtggaattcatctgggataataaacacatattttaa 4109  
Db 3809 aaccttaatgtcttcttcttggaatgtggaattcatctgggataataaacacatattttaa 3868  
QY 4110 cccctcaatattgtctgatacatccgttgcaccccaactcatatagatctgcagcac 4169  
Db 3869 cccctcaatattgtctgatacatccgttgcaccccaactcatatagatctgcagcac 3928  
QY 4170 tcttcgaatggaattgaatggcctgtgattaaataglttgacagatgcattggaatgga 4229  
Db 3929 tcttcgaatggaattgaatggcctgtgattaaataglttgacagatgcattggaatgga 3988  
QY 4230 gattgaagcaatatccagatgcacagattactgttcaacttcaacttcaactatgtttgc 4289  
Db 3989 gattgaagcaatatccagatgcacagattactgttcaacttcaacttcaactatgtttgc 4048  
QY 4290 caccgtgtcccttcaaaaagctcgaactcacccccaagggaggaatgaccccgagacc 4349  
Db 4049 caccgtgtcccttcaaaaagctcgaactcacccccaagggaggaatgaccccgagacc 4108  
QY 4350 tcaagtgaataatccaaaagatggtcgtgcaagtggacttccagaagaacatgaagtcaac 4409  
Db 4109 tcaagtgaataatccaaaagatggtcgtgcaagtggacttccagaagaacatgaagtcaac 4168  
QY 4410 aggaagtaactaccacagggagtaaaatctcgtctaccagaatgtatgtgaagagttcct 4469  
Db 4169 aggaagtaactaccacagggagtaaaatctcgtctaccagaatgtatgtgaagagttcct 4228  
QY 4470 catctccagcagtcagaatgagtcacatgagtcctcttcttccagaatggaagaatlaa 4529  
Db 4229 catctccagcagtcagaatgagtcacatgagtcctcttcttccagaatggaagaatlaa 4288  
QY 4530 ggttttccagggaataccaagactccttcaacacctgtgtgtaactctctagaccacagtt 4589  
Db 4289 ggttttccagggaataccaagactccttcaacacctgtgtgtaactctctagaccacagtt 4248  
QY 4590 actgactcgttactccttcgaattcccccaagattgggtgacccaagattgacctggagat 4649  
Db 4349 actgactcgttactccttcgaattcccccaagattgggtgacccaagattgacctggagat 4408  
QY 4650 ggaagttctgtgctgcagagcacaggaacctctactgactcagcagagatc 4699  
Db 4409 ggaagttctgtgctgcagagcacaggaacctctactgactgaggtgacctgc 4458  
  
RESULT 9  
AAA49232  
ID AAA49232 standard; DNA; 12445 BP.  
XX  
AC AAA49232;  
XX  
DT 26-SEP-2000 (first entry)  
XX  
DE Vector HSReneo for transforming endothelial cells.  
XX  
KW Hemostatic; antianemic; antidiabetic; nootropic; neuroprotective; vector;  
KW osteopathic; antiskinking; immunostimulant; gene therapy; collage;  
KW endothelial cell; peripheral blood; buffy coat cell; VEGF; PCR primer;  
KW vascular endothelial growth factor; bovine brain extract; haemophilia;  
KW Factor VIII; human; transgene; adenosine deaminase deficiency; ss;  
KW sickle cell anaemia; thalassemia; diabetes; alpha-antitrypsin deficiency;  
KW Alzheimer's disease; brain disease; heart disease; immune system defect;  
KW bone fracture; osteoporosis.  
XX  
OS Homo sapiens.  
OS Synthetic.  
PN WO200032750-A1.  
PN  
XX 08-JUN-2000.  
XX

PF 24-NOV-1999; 99WO-US28033.  
XX  
PR 24-NOV-1998; 98US-0109687.  
XX  
XX (MIND ) UNIV MINNESOTA.  
PA (UYEM-) UNIV EMORY.  
PA (HEBB/) HEBBEL R P.  
PA (LIN/) LIN Y.  
XX (LOLL/) LOLLAR J S.  
PI HEBBEL RP, Lin Y, Lollar JS.  
XX  
DR WPI: 2000-412303/35.  
XX  
PT Expanding population of endothelial cells useful to biocompatibilize  
PT implantable medical devices comprises contacting buffy coat cells with  
PT collagen I coated surface in culture medium comprising vascular  
PT endothelial growth factor -  
XX  
XX  
XX Claim 19; Fig 4; 53p; English.  
XX  
XX The invention relates to a method for expanding the population of  
CC endothelial cells (EC) obtained from peripheral blood, by culturing,  
CC in contact with a collagen I coated surface, buffy coat cells obtained  
CC from peripheral mammalian blood in the presence of a culture medium  
CC containing vascular endothelial growth factor (VEGF) and free of bovine  
CC brain extract. EC are useful for treating hemophilia by introducing it  
CC into the blood stream of a mammal, so that an effective amount of  
CC Factor VIII protein is secreted in the blood stream of the mammal.  
CC This sequence represents the vector HSOReNeo where the human factor  
CC VIII gene into which an enhanced green fluorescent protein coding  
CC sequence (H8Q) has been inserted, is subcloned. Transgenic EC transduced  
CC in vitro are useful for improving prosthetic implants. EC is also useful  
CC for diagnosing clotting disorders where indication or disease is  
CC associated with a reduction in the activity of an enzyme. EC is also  
CC useful in gene therapy for treating the variety of diseases including  
CC adenosine deaminase deficiency, sickle cell anemia, thalassemia,  
CC hemophilia, diabetes, alpha-antitrypsin deficiency, brain disorders  
CC such as Alzheimer's disease, heart diseases, defects in immune system,  
CC for repairing bone fractures and to treat or prevent osteoporosis.  
XX  
XX  
SQ Sequence 12445 BP; 3255 A; 2989 C; 2951 G; 3250 T; 0 other;

Query Match 90.6%; Score 4319; DB 21; Length 12445;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 4374; Conservative 0; Mismatches 5; Indels 42; Gaps 1;

QY 266 ccacacatgcaaatagaaactctccacacgctcttctgtgctctttggatctgcttta 325  
DB 763 ccacacatgcaaatagaaactctccacacgctcttctgtgctctttggatctgcttta 822  
QY 326 gtccacacagaagaatactactcctgtgagtggaactgtcattggagacattgtcaaatg 385  
DB 823 gtgcacacagaagaatactactcctgtgagtggaactgtcattggagacattgtcaaatg 882  
QY 386 atctcgggtgagctgtgtggaacgaagattcctccttagagtgccaaaatcctttccat 445  
DB 883 atctcgggtgagctgtgtggaacgaagattcctccttagagtgccaaaatcctttccat 942  
QY 446 tcaaacacctagtcgtgtacaaaaaagactcgtttgttgaattcaaggtttcaacttttca 505  
DB 943 tcaaacacctagtcgtgtacaaaaaagactcgtttgttgaattcaaggtttcaacttttca 1002  
QY 506 acatcgtgaagccaaagccacccctgagtggtctgctagtgccaccatccaggtgag 565  
DB 1003 acatcgtgaagccaaagccacccctgagtggtctgctagtgccaccatccaggtgag 1062  
QY 566 tttaattgaagtggtatattacttaagaacatggtctccatctcttcagcttcaatg 625  
DB 1063 tttaattgaagtggtatattacttaagaacatggtctccatctcttcagcttcaatg 1122  
QY 626 ctgtgtgtatatacctactgtgaagactctctgaggaagctgaatatgatacgaaccagtc 685

DB 1123 ctgtgtgtatatacctactgtgaagactctctgaggaagctgaatatgatacgaaccagtc 1182  
QY 686 aaagggaagaagaagatgatataaagcttccctgtgtggaagccatacatatgtctgcagg 745  
DB 1183 aaagggaagaagaagatgatataaagcttccctgtgtggaagccatacatatgtctgcagg 1242  
QY 746 tccctgaagaagaatggtcccaatggcctctgacccacgtgccttactactcatatcctt 805  
DB 1243 tccctgaagaagaatggtcccaatggcctctgacccacgtgccttactactcatatcctt 1302  
QY 806 ctcatgtgacctgtgtaaagaacttgtaattcgaagcctcatgtgagccctactagatgta 865  
DB 1303 ctcatgtgacctgtgtaaagaacttgtaattcgaagcctcatgtgagccctactagatgta 1362  
QY 866 gagaaggagatctgtgcgaagaagaagaacagacctgtgcgaattatctactctttg 925  
DB 1363 gagaaggagatctgtgcgaagaagaagaacagacctgtgcgaattatctactctttg 1422  
QY 926 ctgtatttgatgaagggaagaagttgtgacatcagaacaagaactccttgatgcagata 985  
DB 1423 ctgtatttgatgaagggaagaagttgtgacatcagaacaagaactccttgatgcagata 1482  
QY 986 gggatgtgcaatctgtctgcggctgtgccttaaatgacacacagtcataatgtgtatgtaaca 1045  
DB 1483 gggatgtgcaatctgtctgcggctgtgccttaaatgacacacagtcataatgtgtatgtaaca 1542  
QY 1046 ggtctctgcaggctgtgtgtgagtcgcaagaagaatcagctcatctgtgcagtatgtaa 1105  
DB 1543 ggtctctgcaggctgtgtgtgagtcgcaagaagaatcagctcatctgtgcagtatgtaa 1602  
QY 1106 tgggacacacatcctcgtgaagtgcactcaatatctcgtgaaggttcacacattctgtgagga 1165  
DB 1603 tgggacacacatcctcgtgaagtgcactcaatatctcgtgaaggttcacacattctgtgagga 1662  
QY 1166 accatcgcgaagcgtctctgtgaaatctcgcgaataaacttctcctactgcgtccaacactct 1225  
DB 1663 accatcgcgaagcgtctctgtgaaatctcgcgaataaacttctcctactgcgtccaacactct 1722  
QY 1226 tgatggaacctgtggaacgtttctactgtttgttcatactctctccacacatgatagga 1285  
DB 1723 tgatggaacctgtggaacgtttctactgtttgttcatactctctccacacatgatagga 1782  
QY 1286 tggaaagcttatgttcaaaagtagacagctgtccgaagaagaaccacactaagaatgaanaata 1345  
DB 1783 tggaaagcttatgttcaaaagtagacagctgtccgaagaagaaccacactaagaatgaanaata 1842  
QY 1346 atgaaagaagcggaaagactatgatataatcttactgattctgaaatgagatgtgtgaagt 1405  
DB 1843 atgaaagaagcggaaagactatgatataatcttactgattctgaaatgagatgtgtgaagt 1902  
QY 1406 ttgatgatgaacaactctcctctccttaccacaaatctgcctcagttgcgaagaagacatccta 1465  
DB 1903 ttgatgatgaacaactctcctctccttaccacaaatctgcctcagttgcgaagaagacatccta 1962  
QY 1466 aaacttgggtacatcaatctgtcgtcgaagaagaagagactgtggaactatgctcccttagtc 1525  
DB 1963 aaacttgggtacatcaatctgtcgtcgaagaagaagagactgtggaactatgctcccttagtc 2022  
QY 1526 tcggcccccagatgacagaagttataaagttcaatatattggaacaatggtgcccacaggaattg 1585  
DB 2023 tcggcccccagatgacagaagttataaagttcaatatattggaacaatggtgcccacaggaattg 2082  
QY 1586 gtgagaagtacaaaaaagctccgatattgtgcatacacagaatggaacacttaagaactcgtgt 1645  
DB 2083 gtgagaagtacaaaaaagctccgatattgtgcatacacagaatggaacacttaagaactcgtgt 2142  
QY 1646 aagctatcagcatgaaatcgaagaatcttgggaaccttacttattggggaagtgtgagaca 1705  
DB 2143 aagctatcagcatgaaatcgaagaatcttgggaaccttacttattggggaagtgtgagaca 2202  
QY 1706 caactgtgatataatttaagaatacaagcagacacatatataacttaccctccagga 1765

Dh 2203 cactgttgatataataaagaatcaagaacgaacataacatcaccctcagaa 2262  
QY 1766 tcactgagctgcgtcccttgatccaagagatccaanaagtgtaaacattgag 1855  
Dh 2263 tcacgagctccgctcttgatccaagagatccaanaagtgtaaacattgag 2322  
QY 1826 atttccaattcgcaggaanaaataccaataaataaagacagtgactgaaag 1885  
Dh 2233 atttccaattcgcaggaanaaataccaataaataaagacagtgactgaaag 2382  
QY 1886 ggcacaactaactcagatccctgctgacgcgtatctactagttcgttaattg 1945  
Dh 2383 ggcacaactaactcagatccctgctgacgcgtatctactagttcgttaattg 2442  
QY 1946 agagaatctagcttcagagatccttgccctccatctgcttaagaagattag 2005  
Dh 2443 agagaatctagcttcagagatccttgccctccatctgcttaagaagattag 2502  
QY 2006 atcaagaaggaacacataatgacagaagaatgacatcgtttctgtattg 2065  
Dh 2503 atcaagaaggaacacataatgacagaagaatgacatcgtttctgtattg 2562  
QY 2066 atgagaacccgaagctggtacccaagaataatacaagcttctcccaatcagctg 2125  
Dh 2563 atgagaacccgaagctggtacccaagaataatacaagcttctcccaatcagctg 2622  
QY 2126 gagtgcagcttgagatccagagttccaagctcccaatcaatgacagcataagct 2185  
Dh 2623 gagtgcagcttgagatccagagttccaagctcccaatcaatgacagcataagct 2682  
QY 2186 atgttttgatagtttgcagtgtcagtttgcttgcataagtgacagtgacattc 2245  
Dh 2683 atgttttgatagtttgcagtgtcagtttgcttgcataagtgacagtgacattc 2742  
QY 2246 taagcatatggagacagactgactctctctgctctctctcgtgatacctccaac 2305  
Dh 2743 taagcatatggagacagactgactctctctgctctctctcgtgatacctccaac 2802  
QY 2306 acaaatggtctatgaagaacacatcccatcccatctcagggagaactgtctca 2365  
Dh 2803 acaaatggtctatgaagaacacatcccatcccatctcagggagaactgtctca 2862  
QY 2366 tgttgatgtaaaacccaggtctatgatactctggtgtgcacaactcagacttgcagaa 2425  
Dh 2863 tgttgatgtaaaacccaggtctatgatactctggtgtgcacaactcagacttgcagaa 2922  
QY 2426 gaggcatagcgccttaactgaagttctctagttgtgaacaagaacatggtatatacg 2485  
Dh 2923 gaggcatagcgccttaactgaagttctctagttgtgaacaagaacatggtatatacg 2982  
QY 2486 aggcacattatgaagaatattcaagcattcgtgagttaaaaacaatgccaattgaacaa 2545  
Dh 2983 aggcacattatgaagaatattcaagcattcgtgagttaaaaacaatgccaattgaacaa 3042  
QY 2546 gaactctctcccaattccaagacacactagcactagcacaagaacattaatgccaacc 2605  
Dh 3043 ggaactctctcccaattccaagacacactagcactagcacaagaacattaatgccaacc 3060  
QY 2606 caccagctctgnaaagcccatcaacgaggaaataactcgtactactctcagtcagatacag 2665  
Dh 3061 caccagctctgnaaagcccatcaacgaggaaataactcgtactactctcagtcagatacag 3120  
QY 2666 aggaatattgactatgataccatcatcagttgaaatgaagaagaattttgaacattc 2725  
Dh 3121 aggaatattgactatgataccatcatcagttgaaatgaagaagaattttgaacattc 3180  
QY 2726 atgataggatgnaaatcagaagcccgcaagcttccaagaagaacgcagacatttta 2785  
Dh 3181 atgataggatgnaaatcagaagcccgcaagcttccaagaagaacgcagacatttta 3240  
QY 2786 ttgctgagtgagagagctctgagattatggatgagtaagctcccaatggttctaagaa 2845  
Dh 3241 ttgctgagtgagagagctctgagattatggatgagtaagctcccaatggttctaagaa 3300

QY 2846 acagggctcgaagtggtgagtgctccctcagttccaagaagaattgtttccagaatattag 2905  
Dh 3301 acagggctcgaagtggtgagtgctccctcagttccaagaagaattgtttccagaatattag 3360  
QY 2906 atggtcccttactcaagcccttaacacgttggaagactaaatgaacatttggactcctg 2965  
Dh 3361 atggtcccttactcaagcccttaacacgttggaagactaaatgaacatttggactcctg 3420  
QY 2966 ggcacataataagagcagaagtgaagaataatcatcgtgtaactttcaagaatacagcct 3025  
Dh 3421 ggcacataataagagcagaagtgaagaataatcatcgtgtaactttcaagaatacagcct 3480  
QY 3026 ctggtccctattccttctatctagcttatttttttgaagaagatcagaaggagaag 3085  
Dh 3481 ctggtccctattccttctatctagcttatttttttgaagaagatcagaaggagaag 3540  
QY 3086 cagaacctagaanaaactttgtcaagcctaatgaacacaaacttacttttgaagaatg 3145  
Dh 3541 cagaacctagaanaaactttgtcaagcctaatgaacacaaacttacttttgaagaatg 3600  
QY 3146 aacataatgagcccaactaaagatgagtttgactgcaaaagcctgggttattctctg 3205  
Dh 3601 aacataatgagcccaactaaagatgagtttgactgcaaaagcctgggttattctctg 3660  
QY 3206 atgttgacctggaanaagatgtgacatcagcctgatttgagccctctgctgcaca 3265  
Dh 3661 atgttgacctggaanaagatgtgacatcagcctgatttgagccctctgctgcaca 3720  
QY 3266 ctacaacactgaacccctgctatggaagacagtgacagtgacaagaaatttctcgtttt 3325  
Dh 3721 ctacaacactgaacccctgctatggaagacagtgacagtgacaagaaatttctcgtttt 3780  
QY 3326 tcaccatcttgatgagacccaagaagctggttaacttcactgaaataatggaagaactgca 3385  
Dh 3781 tcaccatcttgatgagacccaagaagctggttaacttcactgaaataatggaagaactgca 3840  
QY 3386 gggctccctggaatatccagaatggaagatcccatctttaaagaataatctgcctcatg 3445  
Dh 3841 gggctccctggaatatccagaatggaagatcccatctttaaagaataatctgcctcatg 3900  
QY 3446 caatcaatggtctataatagatgatacactaccctggttagtaatggtcagagatcaagaa 3505  
Dh 3901 caatcaatggtctataatagatgatacactaccctggttagtaatggtcagagatcaagaa 3960  
QY 3506 ttcgatggtatctgctcagcaatgggcagcaatgaaacatccattcattcattcagtg 3565  
Dh 3961 ttcgatggtatctgctcagcaatgggcagcaatgaaacatccattcattcattcagtg 4020  
QY 3566 gacatggttccactgctacgnaaaaaaagagatataaattggaactgtataactctatc 3625  
Dh 4021 gacatggttccactgctacgnaaaaaaagagatataaattggaactgtataactctatc 4080  
QY 3626 cagggtcttttgagaacagtgnaaatgttaccatccaagaagctggaatttggcgggtggaat 3685  
Dh 4081 cagggtcttttgagaacagtgnaaatgttaccatccaagaagctggaatttggcgggtggaat 4140  
QY 3686 gacttattgocgagcatctacatgcttggaatgagcacacttttctggtgtacagcaata 3745  
Dh 4141 gacttattgocgagcatctacatgcttggaatgagcacacttttctggtgtacagcaata 4200  
QY 3746 agtgtcagaccccttggaatgagcttctggaacatgaagaatttcagattacaactt 3805  
Dh 4201 agtgtcagaccccttggaatgagcttctggaacatgaagaatttcagattacaactt 4260  
QY 3806 caggacacataatgagcagtgggcccaagaagcttggaacactcatctatctcggatcaatca 3865  
Dh 4261 caggacacataatgagcagtgggcccaagaagcttggaacactcatctatctcggatcaatca 4320  
QY 3866 atgcttgagacccaagaagcccttctctggaatcaagtgatgctgtgtgcaccaatga 3925  
Dh 4321 atgcttgagacccaagaagcccttctctggaatcaagtgatgctgtgtgcaccaatga 4380

|          |             |  |      |
|----------|-------------|--|------|
| OY       | 3926        | ttatcaaggcatcaaacaccagaagttcccgctcagaagttctccagcctctacatcttc             | 3985 |
| Db       | 4381        | ttatcaaggcatcaaacaccagaagttcccgctcagaagttctccagcctctacatcttc             | 4440 |
| OY       | 3986        | agtttcaacatgtaagctcttatgtggaagatgtgcagactatctaggaattcca                  | 4045 |
| Db       | 4441        | agtttcaacatgtaagctcttatgtggaagatgtgcagactatctaggaattcca                  | 4500 |
| OY       | 4046        | cgtgaaccttaatgtgtctcttttggcaatgtgtgatalcaatcvtgggataaacaatattt           | 4105 |
| Db       | 4501        | cgtgaaccttaatgtgtctcttttggcaatgtgtgatalcaatcvtgggataaacaatattt           | 4560 |
| OY       | 4106        | ttaaacctccaatatattgtctcgataacatccgcttggcaaccacatratatgatctgca            | 4165 |
| Db       | 4561        | ttaaacctccaatatattgtctcgataacatccgcttggcaaccacatratatgatctgca            | 4620 |
| OY       | 4166        | gcaccttcgcaggaattgatgtggcgtgtgatttaaatagttgcagcatgtccattgggaa            | 4225 |
| Db       | 4621        | gcaccttcgcaggaattgatgtggcgtgtgatttaaatagttgcagcatgtccattgggaa            | 4680 |
| OY       | 4226        | tggagaagtaagaacataatcagatgcagacagatctatgtcttcaaccttcaacaatagt            | 4285 |
| Db       | 4681        | tggagaagtaagaacataatcagatgcagacagatctatgtcttcaaccttcaacaatagt            | 4740 |
| OY       | 4286        | ttggcaacctgtgtctccttccaagaagctgcactccaacctccaaggaggaagtatgtcctga         | 4345 |
| Db       | 4741        | ttggcaacctgtgtctccttccaagaagctgcactccaacctccaaggaggaagtatgtcctga         | 4800 |
| OY       | 4346        | ggacctaaagttaaatcccaaaagatgtgcctgaagtgtgacttccagaagaacatgaag             | 4405 |
| Db       | 4801        | ggacctaaagttaaatcccaaaagatgtgcctgaagtgtgacttccagaagaacatgaag             | 4860 |
| OY       | 4406        | tccacaggaatcaatctactccaaggagtaaatctctcgtctaccagcatgtatgtgaagagt          | 4465 |
| Db       | 4861        | tccacaggaatcaatctactccaaggagtaaatctctcgtctaccagcatgtatgtgaagagt          | 4920 |
| OY       | 4466        | tccctcaatccagcagatccaagatgtggcatcaagtgtgaactctctttttccaagaatggcaag       | 4525 |
| Db       | 4921        | tccctcaatccagcagatccaagatgtggcatcaagtgtgaactctctttttccaagaatggcaag       | 4980 |
| OY       | 4526        | ttaaagtttttcaggggaataccaagacacctctacacacctgtgtgaactctctagaaccac          | 4585 |
| Db       | 4981        | ttaaagtttttcaggggaataccaagacacctctacacacctgtgtgaactctctagaaccac          | 5040 |
| OY       | 4586        | cgttaactgactcgctacaccttcgaattcaaccocccagagtttggctgcacagattgcctga         | 4645 |
| Db       | 5041        | cgttaactgactcgctacaccttcgaattcaaccocccagagtttggctgcacagattgcctga         | 5100 |
| OY       | 4646        | ggaatggaagttcttggctcgcgaagcacagaacctctactga                              | 4686 |
| Db       | 5101        | ggaatggaagttcttggctcgcgaagcacagaacctctactga                              | 5141 |
| RESULT   | 10          |  |      |
| AAD00121 |             |  |      |
| ID       | AAD00121    | standard; DNA; 11933 BP.   |      |
| AC       | AAD00121;   |  |      |
| XX       |             |  |      |
| DT       | 31-JUL-2000 | (first entry)  |      |
| DE       |             | Recombinant adeno associated vector construct, PAAV-P8-1.                |      |
| XX       |             |  |      |
| KW       |             | Recombinant Adeno Associated Vector; rAAV; PAAV-P8-1; human Factor VIII; |      |
| KW       |             | hepVIT; hNF-3 albumin promoter; human elongation factor-1alpha; ER1alpha |      |
| KW       |             | human growth hormone; hGH; inverted terminal repeat; ITR; haemophilia;   |      |
| XX       |             | gene therapy; ds.  |      |
| XX       |             |  |      |
| OS       |             | Adeno associated virus.  |      |
| XX       |             |  |      |
| PM       |             | WO200023116-A1.  |      |
| PD       |             | 27-APR-2000.   |      |

|    |    |  |   |
|----|----|--|---|
| XX | PF | 19-OCT-1999;   | 98WO-US24495.                             |
| XX | PR | 20-OCT-1998;   | 98US-0104994.                             |
| XX | PR | 24-MAR-1999;   | 99US-0125974.                             |
| XX | PR | 30-JUL-1999;   | 99US-0364862.                             |
| XX | PA | (AVIG-) AVIGEN INC.  |   |
| XX | PI | Couto LB, Colosi PC;   |   |
| XX | DR | WPI: 2000-339536/29.   |   |
| XX | PT | New recombinant adenovirus-comprises at least a portion of Factor VIII     |   |
| XX | PT | operably linked to control sequence -                                      |   |
| XX | PS | Example 2; Fig 5; 92pp; English.   |   |
| CC | XX | The present DNA sequence is a recombinant adenovirus-associated vector,    |   |
| CC | CC | (rAAV) construct, pAAV-F8-1. This expression vector comprises the HNF-3    |   |
| CC | CC | albumin promoter, the first intron (-573 to +985) of human elongation      |   |
| CC | CC | factor-1alpha (Erf1alpha) gene, human Factor VIII coding sequence (hFVIII) |   |
| CC | CC | and polyadenylation signal from human growth hormone (hGH). This sequence  |   |
| CC | CC | is inserted between the AAV inverted terminal repeat (ITR) regions. The    |   |
| CC | CC | hFVIII coding region comprises the heavy chain gene segment with the       |   |
| CC | CC | first 57 bp encoding the 19 amino acid signal peptide, A1 and A2 domains   |   |
| CC | CC | and 5 amino acids from the N-terminus of the B domain. The light chain     |   |
| CC | CC | segment comprises the C-terminal 85 amino acids of B domain and the A3,    |   |
| CC | CC | C1 and C2 domains. Both the heavy and light chain segments are cloned      |   |
| CC | CC | into the same plasmid separated by 42 nucleotides coding for 14 residues   |   |
| CC | CC | of the B domain, that is deleted. This plasmid is operably linked to       |   |
| CC | CC | control sequences, that directs the transcription and translation of the   |   |
| CC | CC | Factor VIII gene. The adeno-associated viral vectors are used for gene     |   |
| CC | CC | therapy to treat haemophilia. This method allows prolonged expression of   |   |
| CC | CC | therapeutic levels of Factor VIII in vivo. The rAAV are used for gene      |   |
| CC | CC | therapy, because of their broad host range, safety profile and duration    |   |
| CC | CC | of expression in the infected hosts.                                       |   |
| CC | XX | Sequence 11933 BP; 3258 A; 2818 C; 2717 G; 3140 T; 0 other:                |   |
| SQ |    |  |   |
|    |    | Query Match  | 89.4%; Score 4259.4; DB 21; Length 11933; |
|    |    | Best Local Similarity  | 98.2%; Pred. No. 0;                       |
|    |    | Matches 4348; Conservative   | 0; Mismatches 1; Indels 78; Gaps 1;       |
| OY |    | 266 ccacatcgtaataagctctccactcgtcttcttcgttgctttgtagattcgttta                | 325                                       |
| DB |    | 422 caccatgcgaataagctctccacctcgtcttcttcgttgctttgtagattcgttta               | 481                                       |
| OY |    | 326' gfgccaccagaagatctactcctgggtgcaaggaaacttcatatgagcataatg                | 385                                       |
| DB |    | 482 gtgccaccagaagatctactcctgggtgcaaggaaacttcatatgagcataatg                 | 541                                       |
| OY |    | 386 attctggtagctgctgtgtgacgcgaagaattctctcctagagtgcgaatatcttccat            | 445                                       |
| DB |    | 542 attctggtagctgctgtgtgacgcgaagaattctctcctagagtgcgaatatcttccat            | 601                                       |
| OY |    | 446 tcaaacactcagtgctgttacaaaaaagactcgttgtatgaattcaaggttcaacctttca          | 505                                       |
| DB |    | 602 tcaaacactcagtgctgttacaaaaaagactcgttgtatgaattcaaggttcaacctttca          | 661                                       |
| OY |    | 506 acatgctgaagcgaagggcacaccttggatgggtctgctagagtcctaccatcaggtcagg          | 565                                       |
| DB |    | 662 acatgctgaagcgaagggcacaccttggatgggtctgctagagtcctaccatcaggtcagg          | 721                                       |
| OY |    | 566 ttatgatatacagtgatctatcaactaatgaacatggtctccatcctctgactctcatg            | 625                                       |
| DB |    | 722 ttatgatatacagtgatctatcaactaatgaacatggtctccatcctctgactctcatg            | 781                                       |
| OY |    | 626 ctgtgtgtgtalccctactctgnaaagctctctggaggagctcgaatatgtatcagaacagtc        | 685                                       |
| DB |    | 782 ctgtgtgtgtalccctactctgnaaagctctctggaggagctcgaatatgtatcagaacagtc        | 841                                       |

|    |      |   |      |
|----|------|---|------|
| OY | 686  | aaagggcgaagaatgatataaagttccttcgtgtggaagccatacatatgtctgcgaag       | 745  |
| Dp | 842  | aaagggcgaagaagaatgatataaagttccttcgtgtggaagccatacatatgtctgcgaag    | 901  |
| OY | 746  | tcctcgaaagaatgtgtcccatctgccttcgaccactgtgccttaaccatacatctt         | 805  |
| Dp | 902  | tcctcgaaagaatgtgtcccatctgccttcgaccactgtgccttaaccatacatctt         | 961  |
| OY | 806  | ctcatctgcgaacctgtgtataaagaacttgaattcagcgctcatcttgagaccctactagatga | 865  |
| Dp | 962  | ctcatctgcgaacctgtgtataaagaacttgaattcagcgctcatcttgagaccctactagatga | 1021 |
| OY | 866  | gagagaaggaaatctgtgcgcgaagaaaagacacagaccttgcacaataattactacttttg    | 925  |
| Dp | 1022 | gagagaaggaaatctgtgcgcgaagaaaagacacagaccttgcacaataattactacttttg    | 1081 |
| OY | 926  | ctgtattctga tgaaggggaaaagtctgcacatcgagaacaaagaacctcttgatgcagagta  | 985  |
| Dp | 1082 | ctgtattctga tgaaggggaaaagtctgcacatcgagaacaaagaacctcttgatgcagagta  | 1141 |
| OY | 986  | gggaatctgcacatctctctgcggcctgtgcctaaatgcaacagtgcaatggttatgtaaca    | 1045 |
| Dp | 1142 | gggaatctgcacatctctctgcggcctgtgcctaaatgcaacagtgcaatggttatgtaaca    | 1201 |
| OY | 1046 | ggtctctgcgaagttctgtattgtgga tgcacagagaaaatcgaattcattgtgagatgttgaa | 1105 |
| Dp | 1202 | ggtctctgcgaagttctgtattgtgga tgcacagagaaaatcgaattcattgtgagatgttgaa | 1261 |
| OY | 1106 | tgggcaaccaactcctctgaagtgtcacataatacttcctgaaagtacacatcttctgtgaaga  | 1165 |
| Dp | 1262 | tgggcaaccaactcctctgaagtgtcacataatacttcctgaaagtacacatcttctgtgaaga  | 1321 |
| OY | 1166 | accatctgcgaaggtctctctgtgaaatcttcgcaataacttctctactctcaaacctct      | 1225 |
| Dp | 1322 | accatctgcgaaggtctctctgtgaaatcttcgcaataacttctctactctcaaacactct     | 1381 |
| OY | 1226 | tgatgagaccttggacaagtttctactctgttctgtcatactcttcccaacaacatgatagca   | 1285 |
| Dp | 1382 | tgatgagaccttggacaagtttctactctgttctgtcatactcttcccaacaacatgatagca   | 1441 |
| OY | 1286 | tggaagcttatgtccaagttagacagctgtctcgaaggaaaccccaactacgaatgataaata   | 1345 |
| Dp | 1442 | tggaagcttatgtccaagttagacagctgtctcgaaggaaaccccaactacgaatgataaata   | 1501 |
| OY | 1346 | atgaagaagaaggagaagcctatgatactgacatcttaacgattctgaaatgtgattgtcgaagt | 1405 |
| Dp | 1502 | atgaagaagaaggagaagcctatgatactgacatcttaacgattctgaaatgtgattgtcgaagt | 1561 |
| OY | 1406 | ttgatgtatgacaacctctcctctcttaaccaaatctcgctcagttgtgcagaagaacctccta  | 1465 |
| Dp | 1562 | ttgatgtatgacaacctctcctctcttaaccaaatctcgctcagttgtgcagaagaacctccta  | 1621 |
| OY | 1466 | aaacttggagacatctcatgtgcgtcgtcgaaggagaagactggagacatgtctcccttagacc  | 1525 |
| Dp | 1622 | aaacttggagacatctcatgtgcgtcgtcgaaggagaagactggagacatgtctcccttagacc  | 1681 |
| OY | 1526 | tcgcccccgga tgcacagaagtataaagaagcacaattctgaacatgtgcctcagcgaagtgtg | 1585 |
| Dp | 1682 | tcgcccccgga tgcacagaagtataaagaagcacaattctgaacatgtgcctcagcgaagtgtg | 1741 |
| OY | 1586 | gttgaggaaatcacaaaaagttccgatttatgtgcatacagatgaaacctttaagaactcgtg   | 1645 |
| Dp | 1742 | gttgaggaaatcacaaaaagttccgatttatgtgcatacagatgaaacctttaagaactcgtg   | 1801 |
| OY | 1646 | aaagctatcagcatgataatcaggaatcttggagaccttactatgtgggaagtgtgagaaca    | 1705 |
| Dp | 1802 | aaagctatcagcatgataatcaggaatcttggagaccttactatgtgggaagtgtgagaaca    | 1861 |
| OY | 1706 | caactgttatattatatttaagaatcaagaagcagacagacatalaacaactaacccacgaa    | 1765 |
| Dp | 1862 | caactgttatattatatttaagaatcaagaagcagacagacatalaacaactaacccacgaa    | 1921 |

|    |      |  |       |
|----|------|--|-------|
| QY | 1766 | tcactgtagtccgctcccttgatattcaaggagatlaeccaaagagtgtaaacatttgaag    | 18235 |
| Db | 1922 | tcactgtagtccgctcccttgatattcaaggagatlaeccaaagagtgtaaacatttgaag    | 19811 |
| QY | 1826 | atttccaattctgcaggaggaaattccaatatataatggagctgtagttagaagtg         | 18855 |
| Db | 1962 | atttccaattctgcaggaggaaattccaatatataatggagctgtagttagaagtg         | 20411 |
| QY | 1886 | ggccaactlaaatcagacccctcggtgcctgaccccgctattactagtttcgttaatag      | 19451 |
| Db | 2042 | ggccaactlaaatcagacccctcggtgcctgaccccgctattactagtttcgttaatag      | 21011 |
| QY | 1946 | agagagatctagcttccaggactatgccccttcctcaactcgtctacaagaactctag       | 20051 |
| Db | 2102 | agagagatctagcttccaggactatgccccttcctcaactcgtctacaagaactctag       | 21611 |
| QY | 2006 | atcaaaagggnaaccgaataatgtaagacaaagggaattgaactctgtttctgatttg       | 20655 |
| Db | 2162 | atcaaaagggnaaccgaataatgtaagacaaagggaattgaactctgtttctgatttg       | 22211 |
| QY | 2066 | atggaacccgaagcttggtaacctacccaagaagatataaacgcttctccccaatccagtg    | 21255 |
| Db | 2222 | atggaacccgaagcttggtaacctacccaagaagatataaacgcttctccccaatccagtg    | 22811 |
| QY | 2126 | gagtgcagcttgaagatccagagttccaaagctccaaactcatgacagatcaatggtc       | 21855 |
| Db | 2282 | gagtgcagcttgaagatccagagttccaaagctccaaactcatgacagatcaatggtc       | 23411 |
| QY | 2186 | atgtttcttgatagtttcagttgtcagtttgtttgtaataaggtgtaactgtacatc        | 22451 |
| Db | 2342 | atgtttcttgatagtttcagttgtcagtttgtttgtaataaggtgtaactgtacatc        | 24011 |
| QY | 2246 | taagcatggggacccaagacttcccttcgtctctctccctgataataactccaac          | 23051 |
| Db | 2402 | taagcatggggacccaagacttcccttcgtctctctccctgataataactccaac          | 24611 |
| QY | 2306 | acaaatctgtctatgaagaacacatcaacctatcccatctccagagagaactgcttca       | 23655 |
| Db | 2462 | acaaatctgtctatgaagaacacatcaacctatcccatctccagagagaactgcttca       | 25211 |
| QY | 2366 | tgctgataggaaaacccaaggtctatgattctgggtgtgccaaactccagacttcggaaca    | 24255 |
| Db | 2522 | tgctgataggaaaacccaaggtctatgattctgggtgtgccaaactccagacttcggaaca    | 25811 |
| QY | 2426 | gaggcatgacccgcttactgaaggttctcagttgtgtaagaagacactgtgattatcg       | 24851 |
| Db | 2582 | gaggcatgacccgcttactgaaggttctcagttgtgtaagaagacactgtgattatcg       | 26411 |
| QY | 2486 | aggacagttatgaagatatcttcagcactgctgtgataaaaacaatgcatltgaacca       | 25451 |
| Db | 2642 | aggacagttatgaagatatcttcagcactgctgtgataaaaacaatgcatltgaacca       | 27011 |
| QY | 2546 | gaagcttctccagaattcaagaacacctagcactagggcaaaagcaatttaatgccacc      | 26051 |
| Db | 2702 | gaagcttctccagaattcaagaacacctagcactagggcaaaagcaatttaatgccacc      | 27611 |
| QY | 2606 | caccagctcttgaaacggcatcaacgggaaataactgttactcttctaagtcagatcaag     | 26655 |
| Db | 2762 | caccagctcttgaaacggcatcaacgggaaataactgttactcttctaagtcagatcaag     | 28211 |
| QY | 2666 | aggaaatctgatactatgataccatatccattgtaaaatggaaggaattttgacatt        | 28031 |
| Db | 2744 | aggaaatctgatactatgataccatatccattgtaaaatggaaggaattttgacatt        | 28631 |
| QY | 2746 | atgatgagatgaaatcacagagcccccgaagcttccaagaagaacacgacataattta       | 27851 |
| Db | 2804 | atgatgagatgaaatcacagagcccccgaagcttccaagaagaacacgacataattta       | 28631 |
| QY | 2786 | tttgtctgacgttggaagagctctgtgattatgggataagtagtgcctcccaatgttctaagaa | 28451 |
| Db | 2864 | tttgtctgacgttggaagagctctgtgattatgggataagtagtgcctcccaatgttctaagaa | 29231 |
| QY | 2846 | acaaagctccaagatgagacagctccctccacttccaagaanaattttccaggaatttaactg  | 29051 |



Db 2924 acagggctcagagtgagagtgctccctcagcttcacgaagaagtgtttccaggaatttactg 2983  
QY 2906 atggctcttactacagcccttataccgttgagaaactaaatgaacatttggactccgg 2965  
Db 2984 atggctcttactacagcccttataccgttgagaaactaaatgaacatttggactccgg 3043  
QY 2966 ggcacataataagagcagaagltgaagaataatactgtaacttcagaaatcagacct 3025  
Db 3044 ggcacataataagagcagaagltgaagaataatactgtaacttcagaaatcagacct 3103  
QY 3026 ctggtccctattccttcttattctagccttatttcttgaagaagatcagaggcagaag 3085  
Db 3104 ctggtccctattccttcttattctagccttatttcttgaagaagatcagaggcagaag 3163  
QY 3086 cagaacctagaanaaactttgtcaagcctaaatgaacccaacttactttggaagaatgc 3145  
Db 3164 cagaacctagaanaaactttgtcaagcctaaatgaacccaacttactttggaagaatgc 3223  
QY 3146 aacatcatatgacacccaactaaagaatgaatltgactgcaaaagcctggctattctctg 3205  
Db 3224 aacatcatatgacacccaactaaagaatgaatltgactgcaaaagcctggctattctctg 3283  
QY 3206 atgttgacctggaanaaagatgtgacacagagcctgattggagcccttctgtctgcca 3265  
Db 3284 atgttgacctggaanaaagatgtgacacagagcctgattggagcccttctgtctgcca 3343  
QY 3266 ctacaacactggaacccctgctcatgtgagacaaagtacagatcagaaatttgcctgctt 3325  
Db 3344 ctacaacactggaacccctgctcatgtgagacaaagtacagatcagaaatttgcctgctt 3403  
QY 3326 tcacacacttctgtatgagaccaaagctgtgtacttcactgaaataatggaagaatctga 3385  
Db 3404 tcacacacttctgtatgagaccaaagctgtgtacttcactgaaataatggaagaatctga 3463  
QY 3386 gggtccctgcaataatcagaatggaagatccacttttaagagaatattatgcttccatg 3445  
Db 3464 gggtccctgcaataatcagaatggaagatccacttttaagagaatattatgcttccatg 3523  
QY 3446 caatcaactgctacataatgatacaactctgcttaagtaatgctcagaatcacaaga 3505  
Db 3524 caatcaactgctacataatgatacaactctgcttaagtaatgctcagaatcacaaga 3583  
QY 3506 ttcgaatgatatctgtctcagaatgagcagaatgaataccatctctatctcaattcagtg 3565  
Db 3584 ttcgaatgatatctgtctcagaatgagcagaatgaataccatctctatctcaattcagtg 3643  
QY 3566 gacatggttctacactgtacgaaanaaagagagtaaaatggcactgtgacaactctctc 3625  
Db 3644 gacatggttctacactgtacgaaanaaagagagtaaaatggcactgtgacaactctctc 3703  
QY 3626 caggtgttcttgagacagtggaatgttaccatccaagctggaatttgcggtgtgaat 3685  
Db 3704 caggtgttcttgagacagtggaatgttaccatccaagctggaatttgcggtgtgaat 3763  
QY 3686 ggccttattgagcagatctacatgctgtggaatgagcaacttttctgtgtacagcaata 3745  
Db 3764 ggccttattgagcagatctacatgctgtggaatgagcaacttttctgtgtacagcaata 3823  
QY 3746 agtgtcagactccctctgggaatgtctctgacacataggaatttcaagtttaagctt 3805  
Db 3824 agtgtcagactccctctgggaatgtctctgacacataggaatttcaagtttaagctt 3883  
QY 3806 caggaacataatgacagtgagccccaagctgacagactcatattccgatatcaaca 3865  
Db 3884 caggaacataatgacagtgagccccaagctgacagactcatattccgatatcaaca 3943  
QY 3866 atgccttgagacccaagagaccttctcttgatcaagtgagatcgttggcacaatga 3925  
Db 3944 atgccttgagacccaagagaccttctcttgatcaagtgagatcgttggcacaatga 4003  
QY 3926 ttattcagcatcaagaaccaggtgcccgtcagaagttctcagacctctacatctctc 3985

Db 4004 ttattcagcatcaagaaccaggtgcccgtcagaagttctcagacctctacatctctc 4063  
QY 3986 agttatcatcatgatalagcttctgatggaaagtgagcaacttaccgaatcgaatccaa 4045  
Db 4064 agttatcatcatgatalagcttctgatggaaagtgagcaacttaccgaatcgaatccaa 4123  
QY 4046 ctggaacctaagtgctctcttggaatgtggaatcactctggaataaacaataatctt 4105  
Db 4124 ctggaacctaagtgctctcttggaatgtggaatcactctggaataaacaataatctt 4183  
QY 4106 ttaacctccaattatgtctgatacatccgtttggacccaactcatatagatcttgcga 4165  
Db 4184 ttaacctccaattatgtctgatacatccgtttggacccaactcatatagatcttgcga 4243  
QY 4166 gaactcttcgaatgagtgatgtgagctgtgatttaaatagttgacagatgccattggaa 4225  
Db 4244 gaactcttcgaatgagtgatgtgagctgtgatttaaatagttgacagatgccattggaa 4303  
QY 4226 tggagagtaaaagcaatatcagatgacagatctactgttcaactcttaccataatgt 4285  
Db 4304 tggagagtaaaagcaatatcagatgacagatctactgttcaactcttaccataatgt 4363  
QY 4286 ttggacactgtgtctctcccaaaagctggaacttcaactccaaggagagtaatgacctgga 4345  
Db 4364 ttggacactgtgtctctcccaaaagctggaacttcaactccaaggagagtaatgacctgga 4423  
QY 4346 gacctcaagtgaataatccaaaagtggtctgcaagtggacttccagaagaacatgaaag 4405  
Db 4424 gacctcaagtgaataatccaaaagtggtctgcaagtggacttccagaagaacatgaaag 4483  
QY 4406 tcaacagagtaactactcaatgaagtaaaatctctgcttaccagatgtatgtgaagagt 4465  
Db 4484 tcaacagagtaactactcaatgaagtaaaatctctgcttaccagatgtatgtgaagagt 4543  
QY 4466 tccatattccagcagatgaagaatggcacaatgagtgactctcttcttcagaatggcagaag 4525  
Db 4544 tccatattccagcagatgaagaatggcacaatgagtgactctcttcttcagaatggcagaag 4603  
QY 4526 taaaggttcttcagggaaatcaagactctcacaactgtgtgaactctctagaccac 4585  
Db 4604 taaaggttcttcagggaaatcaagactctcacaactgtgtgaactctctagaccac 4663  
QY 4586 cgttactgactcgttactcttctgaatcaaccccaagatgtggtgacacaaattggcctga 4645  
Db 4664 cgttactgactcgttactcttctgaatcaaccccaagatgtggtgacacaaattggcctga 4723  
QY 4646 ggaatgaggtctgtggtcgtgagacagagacctctactgactgag 4692  
Db 4724 ggaatgaggtctgtggtcgtgagacagagacctctactgactgag 4770

RESULT 11  
AAN90654  
ID AAN90654 standard; DNA; 4275 BP.  
XX  
AC AAN90654:  
XX  
AC  
XX  
D7 26-JUN-1990 (first entry)  
XX  
DE DNA encoding 740 Arg-1649 Glu human Factor VIII.C.  
XX  
KW Human Factor VIII.C; Ad.RE.neo; 740 Arg-1649 Glu Factor VIII.C;  
KW haemophilia A.  
XX  
OS Homo sapiens.  
XX  
XX  
FH Key  
FT CDS 1..4275  
FT /tag= a  
PN EP30968-A.  
PD 15-MAR-1989.



XX 09-SEP-1988: 88EP-0114769.  
XX  
XX 08-APR-1988: 88JP-0085454.  
XX  
XX (KAGA ) CHEMO-SERO-THERAP.  
XX (TEIJ ) TEIJUN LTD.  
XX  
PI Sugiyama T, Masuda K, Tajima Y, Yonemura H;  
DR WPI: 1989-078467/11.  
DR P-PSDB: AAP91165.  
XX  
PT Prodn. of recombinant human Factor-VIII-C -  
PT using animal cells transformed with a vector contg. the gene for  
PT Factor VIII:C and a promoter  
PS Fig 1(1) - 1(13): ; 32pp; English.  
PS  
XX When translated, Arg-740 of the carboxyl terminus of the H chain is  
XX directly bonded by a peptide bond to Glu-1649 of the amino terminus of  
XX L chain. It is used to transform animal cells so that they produce  
XX human Factor VIII:C. A prefd. expression vector is plasmid Ad.RE.neo.  
XX The expression vector has at least one promoter upstream of AAP90654.  
XX The transformants can constantly and continuously produce human Factor  
XX VIII:C in high yield on a commercial scale. The human Factor VIII:C so  
XX produced is considered to corresp. to the smallest species of active and  
XX intact Factor VIII:C molecules in the human blood plasma. It is useful  
XX for treating haemophilia A patients.  
SQ Sequence 4275 BP; 1245 A; 941 C; 945 G; 1144 T; 0 other;

Query Match 87.6%; Score 4173; DB 10; Length 4275;  
Best Local Similarity 98.0%; Pred. No. 0;  
Matches 4270; Conservative 0; Mismatches 5; Indels 84; Gaps 1;

QY 328 gccacacagaagatcactctgtgtgctgagctgacactgtcattgagcattatgcagaatgct 387  
DB 1 gccacacagaagatcactctgtgtgctgagctgacactgtcattgagcattatgcagaatgct 60  
QY 388 ctccgtgagctcgtctgtgagcgaagattcctcctagagtgcccaaatctttccattc 447  
DB 61 ctccgtgagctcgtctgtgagcgaagattcctcctagagtgcccaaatctttccattc 120  
QY 448 aacacctcagtcgtgtacaaaagactcgtttgttagaattcaaggttcaacctttcaac 507  
DB 121 aacacctcagtcgtgtacaaaagactcgtttgttagaattcaaggttcaacctttttcaac 180  
QY 508 atcgctcaagcgaagcgcacctgagtggtctgtctagctgctcattccagctgagtggt 567  
DB 181 atcgctcaagcgaagcgcacctgagtggtctgtctagctgctcattccagctgagtggt 240  
QY 568 tatgatacagtgctcattcaacttaagaacatggtctccattcctgttcagcttctagct 627  
DB 241 tatgatacagtgctcattcaacttaagaacatggtctccattcctgttcagcttctagct 300  
QY 628 gtgtgtgtatcctcacttggaagcttctgagggagctgaatgatgtatcaagacagtcac 687  
DB 301 gtgtgtgtatcctcacttggaagcttctgagggagctgaatgatgtatcaagacagtcac 360  
QY 688 agggagaagaagatgataaagcttctcctgtgtggaagcattatgtctgagcagtc 747  
DB 361 agggagaagaagatgataaagcttctcctgtgtggaagcattatgtctgagcagtc 420  
QY 748 ctgaaagagatggttccaaatggcctctgacccactgtgcttactaccatactcttct 807  
DB 421 ctgaaagagatggttccaaatggcctctgacccactgtgcttactaccatactcttct 480  
QY 808 catgtgagctgtgtaaaagacttgaattcaggcctcatttgagagcctactagtataga 867  
DB 481 catgtgagctgtgtaaaagacttgaattcaggcctcatttgagagcctactagtataga 540

QY 868 gaaggagctcgtgccaaagaaagacacagacttgcacaatttatactacttttgcct 927  
DB 541 gaaggagctcgtgccaaagaaagacacacacttgcacaatttatactacttttgcct 600  
QY 928 gtaattgatgaaggaagaaaglttgacatcagaacaagaactccttgatgcagataga 987  
DB 601 gtaattgatgaaggaagaaaglttgacatcagaacaagaactccttgatgcagataga 660  
QY 988 gatgtgcactcgtcctcgagcctgctctaaatgacacagtcacatgtgttatgtaacag 1047  
DB 661 gatgtgcactcgtcctcgagcctgctctaaatgacacagtcacatgtgttatgtaacag 720  
QY 1048 tctctgcagagtcgtgttgatgacacaggaatcagtcattgagcattgagtgagtg 1107  
DB 721 tctctgcagagtcgtgttgatgacacaggaatcagtcattgagcattgagtgagtgag 780  
QY 1108 gccacacactcctgaaatgacatcaataatctcctgaaaggtacacacatttctgtgaggaac 1167  
DB 781 gccacacactcctgaaatgacatcaataatctcctgaaaggtacacacatttctgtgaggaac 840  
QY 1168 catgcccagagcgtccttggaatctcgcgaataacttctcactcgtcctaacaactctg 1227  
DB 841 catgcccagagcgtccttggaatctcgcgaataacttctcactcgtcctaacaactctg 900  
QY 1228 atggaacttggagaagttcactgcttcttgcataatcttcccaacaatgtagtcag 1287  
DB 901 atggaacttggagaagttcactgcttcttgcataatcttcccaacaatgtagtcag 960  
QY 1288 gaagcttatgtcagaatagacagctgttccagaggaaccccaactaagaaatgaataat 1347  
DB 961 gaagcttatgtcagaatagacagctgttccagaggaaccccaactaagaaatgaataat 1020  
QY 1348 gaagaagcggaaagacatgatgatgtatcttactgattcgtgaatgtagtgcaggttt 1407  
DB 1021 gaagaagcggaaagacatgatgatgtatcttactgattcgtgaatgtagtgcaggttt 1080  
QY 1408 gatgatacaactctcctccttccatccaaatcgctcagttgccaagaagatcctaa 1467  
DB 1081 gatgatacaactctcctccttccatccaaatcgctcagttgccaagaagatcctaa 1140  
QY 1468 acttgggtacatcacattgctgtctgaaagaggaagctggagctatgctccttactcctc 1527  
DB 1141 acttgggtacatcacattgctgtctgaaagaggaagctggagctatgctccttactcctc 1200  
QY 1528 gcccccagatgacagaagtataaaagtcaatatgtgaacaatgcccctcagcgattggt 1587  
DB 1201 gcccccagatgacagaagtataaaagtcaatatgtgaacaatgcccctcagcgattggt 1260  
QY 1588 aggaagtacaanaaagctcgatttatggaatcacagatgaacacctttaagactgtga 1647  
DB 1261 aggaagtacaanaaagctcgatttatggaatcacagatgaacacctttaagactgtga 1320  
QY 1648 gctatcagacatgaatcagaagaaatcttggaaccttacttattatgggaaagtgtgagacaa 1707  
DB 1321 gctatcagacatgaatcagaagaaatcttggaaccttacttattatgggaaagtgtgagacaa 1380  
QY 1708 ctgtgtatataatttaagaatacgaagcagaacacataacatcctcctcaggaatc 1767  
DB 1381 ctgtgtatataatttaagaatacgaagcagaacacataacatcctcctcaggaatc 1440  
QY 1768 actgatgtccgtccttctgtattcaaggagattacaaaaggtgttaaaacttgaagat 1827  
DB 1441 actgatgtccgtccttctgtattcaaggagattacaaaaggtgttaaaacttgaagat 1500  
QY 1828 ttcccaattctgcgaaggaataatcacaataataaaatgagcagtgagctgtagaagaatgg 1887  
DB 1501 ttcccaattctgcgaaggaataatcacaataataaaatgagcagtgagctgtagaagaatgg 1560  
QY 1888 ccaactcaaatcagatcctcgtggtccgacccgctatattactctagtttggttaataagag 1947  
DB 1561 ccaactcaaatcagatcctcgtggtccgacccgctatattactctagtttggttaataagag 1620  
QY 1948 agagatctagcttcagactcattggtcctctcctatcctgtcacaagaatctgtagat 2007

[illegible]

|    |      |   |      |
|----|------|---|------|
| Dh | 2617 | gcccccttccctctctatctctagccttatttcttaagagaaatcaagagcgaaagagca        | 2676 |
| Qy | 3088 | gaaccctaaaaaaaactctgttcaagcctctaaltgaaaccaaactactcttvtgaaagtgcga    | 3147 |
| Dh | 2677 | gaacctagaaaaaaactctgttcaagcctctaaltgaaaccaaactactcttvtgaaagtgcga    | 2736 |
| Qy | 3148 | catcaatavggcaacccctctaaagaatgttgactgtccaagcctvggcttatttctctgat      | 3207 |
| Dh | 2737 | catcaatavggcaacccctctaaagaatgttgactgtccaagcctvggcttatttctctgat      | 2796 |
| Qy | 3208 | gttgaccgvggaaaaagatgtgcactcagcgccgatvtggaccctctctgtctgcgcaact       | 3267 |
| Dh | 2797 | gttgaccgvggaaaaagatgtgcactcagcgccgatvtggaccctctctgtctgcgcaact       | 2856 |
| Qy | 3268 | aacacactgaaacccctgtctcatatvggagacaagtgcacgtacaggaatttgtctgttttc     | 3327 |
| Dh | 2857 | aacacactgaaacccctgtctcatatvggagacaagtgcacgtacaggaatttgtctgttttc     | 2916 |
| Qy | 3328 | accatcttcttgatvggacccaagctgttactctcaactgaaataatvgaaagaactctgcag     | 3387 |
| Dh | 2917 | accatcttcttgatvggacccaagctgttactctcaactgaaataatvgaaagaactctgcag     | 2976 |
| Qy | 3388 | gtccctctgcaatctccagatvggaagatcccaacttttaagagaattatctgtcttcatga      | 3447 |
| Dh | 2977 | gtccctctgcaatctccagatvggaagatcccaacttttaagagaattatctgtcttcatga      | 3036 |
| Qy | 3448 | atcaaattgctatacaatvggatacactaccgtcttgatgaatgtgctccagatccaagaagt     | 3507 |
| Dh | 3037 | atcaaattgctatacaatvggatacactaccgtcttgatgaatgtgctccagatccaagaagt     | 3096 |
| Qy | 3508 | cgatgtgatactgtctcaagaaatvggcaagcaatgaaacaatccatctatcttaattcaagtga   | 3567 |
| Dh | 3097 | cgatgtgatactgtctcaagaaatvggcaagcaatgaaacaatccatctatcttaattcaagtga   | 3156 |
| Qy | 3568 | catgtgttcatctgttcaagaaaaaagagaaatataaaatgtgcactgttaacatctctatcca    | 3627 |
| Dh | 3157 | catgtgttcatctgttcaagaaaaaagagaaatataaaatgtgcactgttatactctatcca      | 3216 |
| Qy | 3628 | ggtgtttttctggaacagvtgaaatatttaccctcccaagcctggaaatttvgcgvggttgaaatgc | 3687 |
| Dh | 3217 | ggtgtttttctggaacagvtgaaatatttaccctcccaagcctggaaatttvgcgvggttgaaatgc | 3276 |
| Qy | 3688 | cttattctgagsgacatctacatgctgtgggaatgagcacaactttctctgtgttgcagcaataag  | 3747 |
| Dh | 3277 | cttattctgagsgacatctacatgctgtgggaatgagcacaactttctctgtgttgcagcaataag  | 3336 |
| Qy | 3748 | tgtaagactcccccgtgggaatgtgcttcctggsgacaataagagatatttcaagatccagcttcca | 3807 |
| Dh | 3337 | tgtaagactcccccgtgggaatgtgcttcctggsgacaataagagatatttcaagatccagcttcca | 3396 |
| Qy | 3808 | ggaacataatgaaacagvtggcccccaagcvtgccaagctcatattatccgatcatcaat        | 3867 |
| Dh | 3397 | ggaacataatgaaacagvtggcccccaagcvtgccaagctcatattatccgatcatcaat        | 3456 |
| Qy | 3868 | gctctgagcaaccaagvgagccctttctcttgatcaagvtgagatcgttgcgccaatagtat      | 3927 |
| Dh | 3457 | gctctgagcaaccaagvgagccctttctcttgatcaagvtgagatcgttgcgccaatagtat      | 3516 |
| Qy | 3928 | attcaagcgatcaagaagaccagvgtgcccgctcagaagtcttccaagcctctaactctctcaag   | 3987 |
| Dh | 3517 | attcaagcgatcaagaagaccagvgtgcccgctcagaagtcttccaagcctctaactctctcaag   | 3576 |
| Qy | 3988 | tttatcatcatgatagtattcttgatvggaaagaatgtgcagacttatacgaagaagaattccact  | 4047 |
| Dh | 3577 | tttatcatcatgatagtattcttgatvggaaagaatgtgcagacttatacgaagaagaattccact  | 3636 |
| Qy | 4048 | ggaacacttaatgvgctctctcttctgcaatgtgtgatactctcgggataaaacacataatttt    | 4107 |
| Dh | 3637 | ggaacacttaatgvgctctctcttctgcaatgtgtgatactctcgggataaaacacataatttt    | 3696 |
| Qy | 4108 | aaacctccaattatctgtctgatacatcogtttgaccccaactccatataagatctgcagc       | 4167 |
| Dh | 3697 | aaacctccaattatctgtctgatacatcogtttgaccccaactccatataagatctgcagc       | 3756 |

OY 4168 actcttcgcatggaggttgatggcgtgtgatttaataatagttgcagcatgccatggaaatg 4227  
|||||  
Db 3757 actcttcgcatggaggttgatggcgtgtgatttaataatagttgcagcatgccatggaaatg 3816  
OY 4228 gagaagtaagaacatatacatgacatgacagattactgtctcatctacttacttacaataatgtt 4287  
|||||  
Db 3817 gagagtaagaacatatacatgacatgacagattactgtctcatctacttacttacaataatgtt 3876  
OY 4288 gccacccgtgtctctcttcaaaaagctgcgacttcaacctccaaggagagatgctgtgaga 4347  
|||||  
Db 3877 gccacccgtgtctctcttcaaaaagctgcgacttcaacctccaaggagagatgctgtgaga 3936  
OY 4348 cctcaggttgaaataccaanaagtgctgcgaagtgcgaactccagaagacaataaagtic 4407  
|||||  
Db 3937 cctcaggttgaaataccaanaagtgctgcgaagtgcgaactccagaagacaataaagtic 3996  
OY 4408 acaagagtaactactcagggagtaaaatctctgtcttaccagcatgtatgtgaagagttc 4467  
3997 acaagagtaactactcagggagtaaaatctctgtcttaccagcatgtatgtgaagagttc 4056  
OY 4468 ctcatctccagcagtcgaatagtgccatcagtgactctcttttttcagaatggcaagta 4527  
|||||  
Db 4057 ctcatctccagcagtcgaatagtgccatcagtgactctcttttttcagaatggcaagta 4116  
OY 4528 aaggttttcagggaaatcaagactccttcaacactgtgtgaaactctctagaccacgcg 4587  
|||||  
Db 4117 aaggttttttcagggaaatcaagactccttcaacactgtgtgaaactctctagaccacgcg 4176  
OY 4588 ttaactgactgcacttcaatcaccaccagaagttgggtgtgcacagatggcccttagg 4647  
|||||  
Db 4177 ttaactgactgcacttcaatcaccaccagaagttgggtgtgcacagatggcccttagg 4236  
OY 4648 atggaagttctgtggctgcgagggcacaagacctctactga 4686  
|||||  
Db 4237 atggaagttctgtggctgcgagggcacaagacctctactga 4275

RESULT 12  
AAN80446  
ID AAN80446 standard; DNA: 4275 BP.  
XX  
AC AAN80446;  
XX  
DT 10-OCT-1990 (first entry)  
XX  
DE Modified factor VIII:C sequence with the R740-E1649 deletion.  
XX  
KM Modified factor VIII:C; haemophilia; procoagulant;  
XX  
KM blood coagulation; RE deletion; ss.  
XX  
OS Homo sapiens.  
XX  
PN WC8800831-A.  
XX  
PD 11-FEB-1988.  
XX  
PF 31-JUL-1987; 87WO-US01814.  
XX  
PR 01-AUG-1986; 86US-0893375.  
XX  
PA (BIOJ ) BIOGEN NV.  
XX  
PI Pasek MP.  
XX  
DR WPI; 1988-049866/07.  
XX  
DR P-PSDB; AAP80267.  
XX  
PT New DNA sequences encoding modified factor VIII:C - with deletion of DNA  
XX encoding maturation polypeptide, useful for high yield transformation.  
XX  
PS Claim 3; Page 44-45-46-47; 97P; English.  
XX

CC The entire sequence encoding the maturation polypeptide of  
CC factor VIII:C is deleted, i.e. Arg 740-Glu 1649.  
CC The full length Factor VIII:C cDNA has two changes with respect to the  
CC published sequence (EPO application 160457):  
CC C to A at Leu 242 and T to C to C to C change at amino acid residue 1880  
CC (Phe to Leu). The product is produced in approx. 20 times higher  
CC yields than previous recombinant produced factor VIII:C and are more  
CC easily purified. The peptide is used for treating haemophilia A, both  
CC acute and prolonged bleeding.  
XX See also AAN80444 and AAN80447.  
XX  
SQ Sequence 4275 BP; 1245 A; 940 C; 946 G; 1144 T; 0 other;  
Query Match 87.5%; Score 4171.6; DB 9; Length 4275;  
Best Local Similarity 98.0%; Pred. No. 0;  
Matches 4268; Conservative 0; Mismatches 4; Indels 84; Gaps 1;  
OY 328 gccacccaagataactactcgtgtgcagtggaactgtcatgtgactatagcaatgat 387  
|||||  
Db 4 gccacccaagataactactcgtgtgcagtggaactgtcatgtgactatagcaatgat 63  
OY 388 ctggtgtagctgtcgtgtgacgcaagaattctctctagagtgccaatctttccattc 447  
64 ctggtgtagctgtcgtgtgacgcaagaattctctctagagtgccaatctttccattc 123  
OY 448 aacacccatcgtgtgtacaaaagactctgttttagaatcaaggttcaacctttcaac 507  
|||||  
Db 124 aacacccatcgtgtgtacaaaagactctgttttagaatcaaggttcaacctttcaac 183  
OY 508 atcgctaaagccaagggcacccttgatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 567  
184 atcgctaaagccaagggcacccttgatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 243  
OY 568 tatgatacagtgatgatacttaacacttaagaacatggtcttccatcctgtcagttcattgct 627  
244 tatgatacagtgatgatacttaacacttaagaacatggtcttccatcctgtcagttcattgct 303  
OY 628 gttgtgtatctctactgtgaaagctcttcgagggagctgaatatgtatgtatcagaccagttcaa 687  
304 gttgtgtatctctactgtgaaagctcttcgagggagctgaatatgtatgtatcagaccagttcaa 363  
OY 688 agggagaagaagaatgataagctcttccctgtgtggaagccaataatgtctgtgcagtic 747  
364 agggagaagaagaatgataagctcttccctgtgtggaagccaataatgtctgtgcagtic 423  
OY 748 ctgaaagaagaatgltccaatggtccttgacccactgtgcttacttactatcttct 807  
424 ctgaaagaagaatgltccaatggtccttgacccactgtgcttacttactatcttct 483  
OY 808 catgtgacctgtgtaaaagacttgaattcaggtccttatgtgagccctactagatgtaga 867  
484 catgtgacctgtgtaaaagacttgaattcaggtccttatgtgagccctactagatgtaga 543  
OY 868 gaaaggagctgtgcgaagaaagaacacagacactgtgcacaatttactacttctgct 927  
544 gaaaggagctgtgcgaagaaagaacacacacactgtgcacaatttactacttctgct 603  
OY 928 gtattgtatgaaggaaagtgtgacctcagaacaaagaactcctgtatgcagatagg 987  
604 gtattgtatgaaggaaagtgtgacctcagaacaaagaactcctgtatgcagatagg 663  
OY 988 gatgtgcattctgtctgtggtcgtgcttaaaatgacacagtcattgttataaacaag 1047  
664 gatgtgcattctgtctgtggtcgtgcttaaaatgacacagtcattgttataaacaag 723  
OY 1048 tctctgccaagctgtgattgtgacagcagaagaatcagtcattgtgagctgtgagatg 1107  
724 tctctgccaagctgtgattgtgacagcagaagaatcagtcattgtgagctgtgagatg 783  
OY 1108 ggcacccatcctgtgaagtgacatcaatatctcctcgaaggtcacacattctctgtgaggaac 1167  
784 ggcacccatcctgtgaagtgacatcaatatctcctcgaaggtcacacattctctgtgaggaac 843

QY 1168 catcgccagcgctccttggaaatctcgcaataacttctctactgtctcaaaactcttg 1227  
|||||  
Db 844 catcgccagcgctccttggaaatctcgcaataacttctctactgtctcaaaactcttg 903  
QY 1228 atggaaccttggaacagttctactgttttgcatactcttcccacaacaatgagcatg 1287  
|||||  
Db 904 atggaaccttggaacagttctactgttttgcatactcttcccacaacaatgagcatg 963  
QY 1288 gaagcttatgtccaagtatgacagctgtctcagaaggaaccccaactcgaaatgaaataat 1347  
|||||  
Db 964 gaagcttatgtccaagtatgacagctgtctcagaaggaaccccaactcgaaatgaaataat 1023  
QY 1348 gaagaacggaagacataatgatgatcttaactgtcttgcgaatgagatgtgttcagggtt 1407  
|||||  
Db 1024 gaagaacggaagacataatgatgatcttaactgtcttgcgaatgagatgtgttcagggtt 1083  
QY 1408 gatgatgaacaactctctctctcttccaatctgcataagttgcagaagaacatccctaa 1467  
|||||  
Db 1084 gatgatgaacaactctctctctcttccaatctgcataagttgcagaagaacatccctaa 1143  
QY 1468 actgtgggtacatactatcatgtctgtcgaagaaggaatggaactgtctcccttaagcttc 1527  
|||||  
Db 1144 actgtgggtacatactatcatgtctgtcgaagaaggaatggaactgtctcccttaagcttc 1203  
QY 1528 gcccccgtatgcagaagttataaaagttcaataattgnaacaatgycgcctcagcggaattgt 1587  
|||||  
Db 1204 gcccccgtatgcagaagttataaaagttcaataattgnaacaatgycgcctcagcggaattgt 1263  
QY 1588 aggaagttacaaaaaagttccgattttatggtcatacacagatgaaacctttaaagctcgtgaa 1647  
|||||  
Db 1264 aggaagttacaaaaaagttccgattttatggtcatacacagatgaaacctttaaagctcgtgaa 1323  
QY 1648 gctatgaagcatgatcatcgagatcttggacaccttacttatgggggaagtgttgagacaa 1707  
|||||  
Db 1324 gctatgaagcatgatcatcgagatcttggacaccttacttatgggggaagtgttgagacaa 1383  
QY 1708 cgtgtgatataatttaagaatcaaaagcagaacgacataataacatctacccacggaatc 1767  
|||||  
Db 1384 cgtgtgatataatttaagaatcaaaagcagaacgacataataacatctacccacggaatc 1443  
QY 1768 actgatgtccgctcctctgtatctcaagagatctacaaaaagtgtaaaacattttgaagat 1827  
|||||  
Db 1444 actgatgtccgctcctctgtatctcaagagatctacaaaaagtgtaaaacattttgaagat 1503  
QY 1828 ttccaattcgcgcaggaagaatactcaaatataaattgacagtgaccgttagaagatgg 1887  
|||||  
Db 1504 ttccaattcgcgcaggaagaatactcaaatataaattgacagtgaccgttagaagatgg 1563  
QY 1888 ccaactaatatcagatccctcggctgcctgcagccgctatctactagtctgttaataatggag 1947  
|||||  
Db 1564 ccaactaatatcagatccctcggctgcctgcagccgctatctactagtctgttaataatggag 1623  
QY 1948 agagatcttagcttcaagatcatcttgccctctccctcatctgcttaaaaaagaactgttaagat 2007  
|||||  
Db 1624 agagatcttagcttcaagatcatcttgccctctccctcatctgcttaaaaaagaactgttaagat 1683  
QY 2008 caaagaagaacaacagataatgttcaagaacaaggaatgtcatccctgttttcgtatcttgat 2067  
|||||  
Db 1684 caaagaagaacaacagataatgttcaagaacaaggaatgtcatccctgttttcgtatcttgat 1743  
QY 2068 gagaacccgaagctgtgtactcctacagagaataatacaacgcttcttcccacatccagctgga 2127  
|||||  
Db 1744 gagaacccgaagctgtgtactcctacagagaataatacaacgcttcttcccacatccagctgga 1803  
QY 2128 gtgcagcttgagatctcagagattccaagcttccaacatcatgacacagcataatvggctat 2187  
|||||  
Db 1804 gtgcagcttgagatctcagagattccaagcttccaacatcatgacacagcataatvggctat 1863  
QY 2188 gtctttgatagttgtcaggtgttcagttgtttgtcatgaggtgcagctactgtgtatctta 2247  
|||||  
Db 1864 gtctttgatagttgtcaggtgttcagttgtttgtcatgaggtgcagctactgtgtatctta 1923

QY 2248 agcatltggagacacagactgacttct 2307  
|||||  
Db 1924 agcatltggagacacagactgacttct 1983  
QY 2308 aaaaatgtcttaagaagaacaactcaactccctattccattctcaaggaagaactgttctatg 2367  
|||||  
Db 1984 aaaaatgtcttaagaagaacaactcaactccctattccattctcaaggaagaactgttctatg 2043  
QY 2368 tcgattggaaaaacccaggtctatggaattctcgggtgtccacaactcaagatcttcggaaga 2427  
|||||  
Db 2044 tcgattggaaaaacccaggtctatggaattctcgggtgtccacaactcaagatcttcggaaga 2103  
QY 2428 ggcacatgcgccttacttgaaagtttctagtgttgacaagaacactgtgtatatactagag 2487  
|||||  
Db 2104 ggcacatgcgccttacttgaaagtttctagtgttgacaagaacactgtgtatatactagag 2163  
QY 2488 gacaaattatgaaagataattcagcatactgtcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2547  
|||||  
Db 2164 gacaaattatgaaagataattcagcatactgtcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2223  
QY 2548 agctctccagaatccaagaatccaagcactagcaactagtcgaaagcaatttaatgcccacca 2607  
|||||  
Db 2224 ----- 2223  
QY 2608 ccagcttggaaacgcatcaacggygaataactcgtactactctcagttcagatcaagag 2667  
|||||  
Db 2224 ----- gaataactcgtactactctcagttcagatcaagag 2259  
QY 2668 gaatttgactatgatgatataccatatacagttgaatgaagaaggaagattttgaattat 2727  
|||||  
Db 2260 gaatttgactatgatgatataccatatacagttgaatgaagaaggaagattttgaattat 2319  
QY 2728 gatgagatltgaaatccagagcccccgcagcttccaagaagaacaacagacatattat 2787  
|||||  
Db 2320 gatgagatltgaaatccagagcccccgcagcttccaagaagaacaacagacatattat 2379  
QY 2788 gctgcagttggaaggtcctgtgattatggaatggaatggaatggaatggaatggaatggaatggaat 2847  
|||||  
Db 2380 gctgcagttggaaggtcctgtgattatggaatggaatggaatggaatggaatggaatggaatggaat 2439  
QY 2848 aggtctcagaggtgaggtgtcctcagttcgaagaagtgtttccaggaattacagat 2907  
|||||  
Db 2440 aggtctcagaggtgaggtgtcctcagttcgaagaagtgtttccaggaattacagat 2499  
QY 2908 ggtcctcttactcagcccttataccgttggaagaactaaatggaacttttggagctcctggg 2967  
|||||  
Db 2500 ggtcctcttactcagcccttataccgttggaagaactaaatggaacttttggagctcctggg 2559  
QY 2968 ccatataagaagcagaagttggaagataataatcatgtgttaacttcagaaatcagcctct 3027  
|||||  
Db 2560 ccatataagaagcagaagttggaagataataatcatgtgttaacttcagaaatcagcctct 2619  
QY 3028 cgtccctatctctctatcttagccttatttcttaatgaggaagtcagagcagaagga 3087  
|||||  
Db 2620 cgtccctatctctctatcttagccttatttcttaatgaggaagtcagagcagaagga 2679  
QY 3088 gaacctagaaaaaacttgttcaagccttaattgaacccaacttacttcttggaaagtgcga 3147  
|||||  
Db 2680 gaacctagaaaaaacttgttcaagccttaattgaacccaacttacttcttggaaagtgcga 2739  
QY 3148 catcatatggaacccaactaaagatgagttgactgtcgaagcctgtggtcctatctctgtat 3207  
|||||  
Db 2740 catcatatggaacccaactaaagatgagttgactgtcgaagcctgtggtcctatctctgtat 2799  
QY 3208 gttgacctggaaaaagatgtgtcactcagcgt 3267  
|||||  
Db 2800 gttgacctggaaaaagatgtgtcactcagcgt 2859  
QY 3268 aacaacatgaaacctgtcattggtgagagaatgtgacatgacaggaattgtcgtgttttc 3327  
|||||  
Db 2860 aacaacatgaaacctgtcattggtgagagaatgtgacatgacaggaattgtcgtgttttc 2919  
QY 3328 accatctttagatggaaccaaagctgtgtacttcaatgaaaaatgtaaaagaactgcag 3387

```

Db 2920 accactttgttgagaccaaagcgtgacttcaactgtaaaatcgtgaagaagtaactgcagg 2979
|
|
|
QY 3388 gctccctgcaatcctcagatggaagatcccaacttttaagaagatctcgttccatgca 3447
|
|
|
Db 2980 gctccctgcaatcctcagatggaagatcccaacttttaagaagatctcgttccatgca 3039
|
|
|
QY 3448 atcaatgctcaatcctcagatggaagatcccaacttttaagaagatctcgttccatgca 3507
|
|
|
Db 3040 atcaatgctcaatcctcagatggaagatcccaacttttaagaagatctcgttccatgca 3099
|
|
|
QY 3508 cgaatgctcaatcctcagatggaagatcccaacttttaagaagatctcgttccatgca 3567
|
|
|
Db 3100 cgaatgctcaatcctcagatggaagatcccaacttttaagaagatctcgttccatgca 3159
|
|
|
QY 3568 catggttctcactgtaagcaaaaaaagagatataaataagcactgtacactctatcca 3627
|
|
|
Db 3160 catggttctcactgtaagcaaaaaaagagatataaataagcactgtacactctatcca 3219
|
|
|
QY 3628 ggtgtttttgagacagtggaatgttaccatccaagctggaatttgccggttggaatgc 3687
|
|
|
Db 3220 ggtgtttttgagacagtggaatgttaccatccaagctggaatttgccggttggaatgc 3279
|
|
|
QY 3688 ctatgtgagacatctacatgctgagatgagacacacttttctggtgtacagcaataag 3747
|
|
|
Db 3280 ctatgtgagacatctacatgctgagatgagacacacttttctggtgtacagcaataag 3339
|
|
|
QY 3748 tgtcagactccccctggagatgctctctggaacattagagaatttcagatcacagctca 3807
|
|
|
Db 3340 tgtcagactccccctggagatgctctctggaacattagagaatttcagatcacagctca 3359
|
|
|
QY 3808 ggaacataatgagacagtgaggcccaagctggccagactctatctccggtacatcaat 3867
|
|
|
Db 3400 ggaacataatgagacagtgaggcccaagctggccagactctatctccggtacatcaat 3459
|
|
|
QY 3868 gctcggagacacaaaggagcccttcttggatcaaggctgctgtgccaatgatt 3927
|
|
|
Db 3460 gctcggagacacaaaggagcccttcttggatcaaggctgctgtgccaatgatt 3519
|
|
|
QY 3928 attcagagcacaagaccagaggtgcccgttcaagaattctccacccctacatctccag 3987
|
|
|
Db 3520 attcagagcacaagaccagaggtgcccgttcaagaattctccacccctacatctccag 3579
|
|
|
QY 3988 ttatcatcatgtaagtcttggatggaagaagtgccagactctatctccggtacatcaat 4047
|
|
|
Db 3580 ttatcatcatgtaagtcttggatggaagaagtgccagactctatctccggtacatcaat 3639
|
|
|
QY 4048 ggaaccttaatgctctcttggcaatgctggaatcattcgtggaataaacaataatttt 4107
|
|
|
Db 3640 ggaaccttaatgctctcttggcaatgctggaatcattcgtggaataaacaataatttt 3659
|
|
|
QY 4108 aacctccaattatgctcgtacatccgttgcacccaactcaattatagatcgcagc 4167
|
|
|
Db 3700 aacctccaattatgctcgtacatccgttgcacccaactcaattatagatcgcagc 3759
|
|
|
QY 4168 accttcgcatgagtgatggctggtgatttaattggaatggaatggaatggaatggaatg 4227
|
|
|
Db 3760 accttcgcatgagtgatggctggtgatttaattggaatggaatggaatggaatggaatg 3819
|
|
|
QY 4228 gagaagtaagaatcagatgagatgagatgagatgagatgagatgagatgagatgagatg 4287
|
|
|
Db 3820 gagaagtaagaatcagatgagatgagatgagatgagatgagatgagatgagatgagatg 3879
|
|
|
QY 4288 ggaacacgtgctccttcaaaaagctgcagactcactccaagggagagatgagatgagatg 4347
|
|
|
Db 3880 ggaacacgtgctccttcaaaaagctgcagactcactccaagggagagatgagatgagatg 3939
|
|
|
QY 4348 cctcaggtggaataatccaagaagtggtgcaagtggaatggaatggaatggaatggaatgga 4407
|
|
|
Db 3940 cctcaggtggaataatccaagaagtggtgcaagtggaatggaatggaatggaatggaatgga 3999
|
|
|
QY 4408 acagagtaactactcagggagtaaaatctctgcttaccagcatgtaatgtaagaagattc 4467
|
|
|

```

```

Db 4000 acagagtaactactcagggagtaaaatctctgcttaccagcatgtatgtgaagattc 4059
|
|
|
QY 4468 ctcatctccagacatcgaagatgagccatcagatgagctccttttccgaatggaagta 4527
|
|
|
Db 4060 ctcatctccagacatcgaagatgagccatcagatgagctccttttccgaatggaagta 4119
|
|
|
QY 4528 aaggttttccagggaaatcagactccttccacactggtggaactcctagaccaccg 4587
|
|
|
Db 4120 aaggttttccagggaaatcagactccttccacactggtggaactccttagaccaccg 4179
|
|
|
QY 4588 ttactgactcgtactcctcgaatccacccagagttggtgcaacagatgacctgagg 4647
|
|
|
Db 4180 ttactgactcgtactcctcgaatccacccagagttggtgcaacagatgacctgagg 4239
|
|
|
QY 4648 atggaagttctggtgctgcgaagcagactctac 4683
|
|
|
Db 4240 atggaagttctggtgctgcgaagcagactctac 4275
|
|
|

```

## RESULT 13

AAN80447  
ID AAN80447 standard; DNA: 4272 BP.

AC AAN80447;  
XX

DT 10-OCR-1990 (first entry)  
XX

DE Modified factor VIII:C sequence with the R740-D1658 deletion.  
XX

KM Modified factor VIII:C; haemophilia; procoagulant;  
KW blood coagulation; RD deletion; ss.  
XX

OS Homo sapiens.  
XX

PN W08800831-A.  
XX

PD 11-FEB-1988.  
XX

PF 31-JUL-1987; 87MO-US01814.  
XX

PR 01-AUG-1986; 86US-0893375.  
XX

PA (BIOJ ) BIOGEN NV.  
XX

PI Pasek MP;  
XX

DR WPI; 1988-049866/07.  
XX

DR P-PSDB; AAP80268.  
XX

PT New DNA sequences encoding modified factor VIII:C - with deletion of DNA  
encoding maturation polypeptide, useful for high yield transformation.  
XX

PS Claim 3; Page 47-48-49-50; 97P; English.  
XX

CC The RD deletion removes the DNA from Ser 741 to Ser 1657. A major part  
of the sequence encoding the maturation polypeptide of  
factor VIII:C is deleted, i.e. Gln 744 - Asp 1563.  
XX

CC The full length Factor VIII:C cDNA has two changes with respect to the  
published sequence (EPO application 160457):  
XX

CC CNG to CTA at Leu 242 and TGC to CTC change at amino acid residue 1880  
(the to Leu). The product is produced in approx. 20 times higher  
XX

CC yields than previous recombinant produced factor VIII:C and are more  
easily purified. The peptide is used for treating haemophilia A, both  
XX

CC acute and prolonged bleeding.  
XX

CC See also AAN80444 and AAN80446.  
XX

SO Sequence 4272 BP; 1243 A; 941 C; 944 G; 1144 T; 0 other;  
XX

Query Match 87.5%; Score 4170; DB 9; Length 4272;  
Best Local Similarity 98.0%; Pred. No. 0;  
Matches 4267; Conservative 0; Mismatches 5; Indels 84; Gaps 1;

|    |      |   |      |
|----|------|---|------|
| OY | 328  | gaaacccaagaagatgatactctccctggtggtgcagtggaacgctgaatggaatggaatgataatgaaagtat    | 387  |
| Db | 1    | ggcaaccagaagatactatacctctgggtgtgcagtggaactgtcaatggaatggaatggaatgataatgaaagtat | 60   |
| OY | 388  | ctcgtgtagctctgcctctgtgaacgcagaagatttctctccctagaatgtgcataatctttccatctc         | 447  |
| Db | 61   | ctcgtgtagctctgcctctgtgaacgcagaagatttctctccctagaatgtgcataatctttccatctc         | 120  |
| OY | 448  | aacaacctcagtcgtgtacaacaaaagaactcgtttgttagaatcaacggttcaacctttccaac             | 507  |
| Db | 121  | aacaacctcagtcgtgtacaacaaaagaactcgtttgttagaatcaacggttcaacctttccaac             | 180  |
| OY | 508  | atcgtctaagcccaagggccaacctgtgatgtgtctgtcgaagttcttaccatccagctgaagtt             | 567  |
| Db | 181  | atcgtctaagcccaagggccaacctgtgatgtgtctgtcgaagttcttaccatccagctgaagtt             | 240  |
| OY | 568  | tatgataacagttgtgcatctatacctaataagaacatgtgtccatctccctgtacgtttcatgct            | 627  |
| Db | 241  | tatgataacagttgtgcatctatacctaataagaacatgtgtccatctccctgtacgtttcatgct            | 300  |
| OY | 628  | gtctgtgtatctactactgtggaagagctctgtgagggagctgataatgtaatcatgacagctcaac           | 687  |
| Db | 301  | gtctgtgtatctactactgtggaagagctctgtgagggagctgataatgtaatcatgacagctcaac           | 360  |
| OY | 688  | agggaagaagaatgataatgataaagctctccctcgtgtggaagcaatacatgtctgtgcagttc             | 747  |
| Db | 361  | agggaagaagaatgataatgataaagctctccctcgtgtggaagcaatacatgtctgtgcagttc             | 420  |
| OY | 748  | ctgaagaagaatgtgtccaatgtgctctgtgacccaagctgtccttaccatactaatcttct                | 807  |
| Db | 421  | ctgaagaagaatgtgtccaatgtgctctgtgacccaagctgtccttaccatactaatcttct                | 480  |
| OY | 808  | catgtgtgacctgtgtataaagaacttbaatctcaagccctcatcttgagccctactagatgtata            | 867  |
| Db | 481  | catgtgtgacctgtgtataaagaacttbaatctcaagccctcatcttgagccctactagatgtata            | 540  |
| OY | 868  | gaaagggaagctctgtgccaaaggaaagaacacagacctgtgcacaattatactacttttgc                | 927  |
| Db | 541  | gaaagggaagctctgtgccaaaggaaagaacacacacccctgtgcacaattatactacttttgc              | 600  |
| OY | 928  | gtatttcaatgaagaaggaaagtgtgtgcactcagaataaagaagaactcttbatcaggaatag              | 987  |
| Db | 601  | gtatttcaatgaagaaggaaagtgtgtgcactcagaataaagaagaactcttbatcaggaatag              | 660  |
| OY | 988  | gattctgtcatctgtcctgcggcctgtgcctctaaatgtgacacagctcaatgtgttatgtataacgg          | 1047 |
| Db | 661  | gattctgtcatctgtcctgcggcctgtgcctctaaatgtgacacagctcaatgtgttatgtataacgg          | 720  |
| OY | 1048 | tctctgtccaggcttgatattgtatgtgcacagaagaatcgaatcttatgtgcattgtatgtgaatg           | 1107 |
| Db | 721  | tctctgtccaggcttgatattgtatgtgcacagaagaatcgaatcttatgtgcattgtatgtgaatg           | 780  |
| OY | 1108 | ggacaacactctgaagtgacactcaatatctctctgaagttcaacatatcttctgttggggaac              | 1167 |
| Db | 781  | ggacaacactctgaagtgacactcaatatctctctgaagttcaacatatcttctgttggggaac              | 840  |
| OY | 1168 | catgcgcgaagcgcttccctgttgaatatctgcacataacttccctactacgtgcacaacactctg            | 1227 |
| Db | 841  | catgcgcgaagcgcttccctgttgaatatctgcacataacttccctactacgtgcacaacactctg            | 900  |
| OY | 1228 | atggaaccttggacagtttctactcgttttgttcatactcttccccaacaatgatgtgcatg                | 1287 |
| Db | 901  | atggaaccttggacagtttctactcgttttgttcatactcttccccaacaatgatgtgcatg                | 960  |
| OY | 1288 | gaagcttatgttcaaatgtagaacagctgttccagaaggaaacccaactcagaaatgaanaataat            | 1347 |
| Db | 961  | gaagcttatgttcaaatgtagaacagctgttccagaaggaaacccaactcagaaatgaanaataat            | 1020 |
| OY | 1348 | gaagaagcgaagaactatgtatgtgaacttctgaacttctatcttgaaatgtatgtgtgcagttt             | 1407 |
| Db | 1021 | gaagaagcgaagaactatgtatgtgaacttctgaacttctatcttgaaatgtatgtgtgtgcagttt           | 1086 |
| OY | 1408 | gattgatgtacaactctcttcccttataccaatctgtgtctgaagttgtccaaagaagatccttaac           | 1467 |

[illegible]



Db 2161 gacagtaataagatatatttcagcatcttgctgagtaaaaaaatgccaattgaaccaaga 2220  
QY 2548 agctctccaggaattcaagaacacctagcacataggaaaaaaagaatttaattgcacacca 2607  
Db 2221 ----- 2220  
QY 2608 ccagctctgaacgcacataccggaataactcgtactactctcttcagtcagatcaagag 2667  
Db 2221 -----gaaataactcgtactactctcttcagtcagatcaagag 2256  
QY 2668 gaattgactatgataaccatacatcagttgaaatgaaaggaagattttgacatttat 2316  
Db 2257 gaattgactatgataaccatacatcagttgaaatgaaaggaagattttgacatttat 2316  
QY 2728 gataagataaataatcagagcccccagagctttcaaaaagaaaacagacacttttatt 2787  
Db 2217 gataagataaataatcagagcccccagagctttcaaaaagaaaacagacacttttatt 2787  
QY 2788 gctgcagctgagagagctctggaattatggaatgagtaagctccccaacagtcttaagaag 2847  
Db 2377 gctgcagctgagagagctctggaattatggaatgagtaagctccccaacagtcttaagaag 2436  
QY 2848 agggctcagaagtgcagctgcctcagttcaagaagaagtgtttccaaggaattactgat 2907  
Db 2437 agggctcagaagtgcagctgcctcagttcaagaagaagtgtttccaaggaattactgat 2496  
QY 2908 ggtcctcttactagcccttatccgttgcgtaagaaataatgaacatttggaacctcgggg 2967  
Db 2497 ggtcctcttactagcccttatccgttgcgtaagaaataatgaacatttggaacctcgggg 2967  
QY 2968 ccataataagacagaagcttgaagaataatcatgtaactcttcagaaatccagccctc 3027  
Db 2557 ccataataagacagaagcttgaagaataatcatgtaactcttcagaaatccagccctc 2616  
QY 3028 cgtccctactcttctatctcagccttattctctatgaaagaaatcaagagcaaggagca 3087  
Db 2617 cgtccctactcttctatctcagccttattctctatgaaagaaatcaagagcaaggagca 2676  
QY 3088 gaacctagaataaactttgtcgaagcctaagtaagaaacaaacttacttttggaaagtgc 3147  
Db 2677 gaacctagaataaactttgtcgaagcctaagtaagaaacaaacttacttttggaaagtgc 2736  
QY 3148 catcataatgacacccaataaagaatgatttgaactgcaaaacccgtggcttatctctgat 3207  
Db 2737 catcataatgacacccaataaagaatgatttgaactgcaaaacccgtggcttatctctgat 2796  
QY 3208 gttgacctgtaaaaaagatgtgcactcagctcagctgattggaccctctgctgcgcacact 3267  
Db 2797 gttgacctgtaaaaaagatgtgcactcagctcagctgattggaccctctgctgcgcacact 2856  
QY 3268 aacacactgaaacctgtctcattgagagaacagtgcagtaagaatttggctcgttttc 3327  
Db 2857 aacacactgaaacctgtctcattgagagaacagtgcagtaagaatttggctcgttttc 2916  
QY 3328 accatctttgatagacccaanaagctgttactcactcgaanaataatggaaaanaacgtcag 3387  
Db 2917 accatctttgatagacccaanaagctgttactcactcgaanaataatggaaaanaacgtcag 2976  
QY 3388 gctccctgcaataatccagatggaagatccacatttaagaagaataatccgtctcagta 3447  
Db 2977 gctccctgcaataatccagatggaagatccacatttaagaagaataatccgtctcagta 3036  
QY 3448 ataataatgacataatgatactacactcgtttagtaatgcttaagaataaagaatt 3507  
Db 3037 ataataatgacataatgatactacactcgtttagtaatgcttaagaataaagaatt 3096  
QY 3508 cgaatgtatctgcagcagctgagcagaatgaataaacaatccatcttatcttcagttga 3567  
Db 3097 cgaatgtatctgcagcagctgagcagaatgaataaacaatccatcttatcttcagttga 3156  
QY 3568 catgttctcagctgacgaataaagaaggtataaataatgacactgtacatctctatcca 3627  
Db 3157 catgttctcagctgacgaataaagaaggtataaataatgacactgtacatctctatcca 3216

QY 3628 gttgtttttgagacagtggaatgtttacacatcccaaaagcttggaatttggcgggtggaatgc 3687  
Db 3217 gttgtttttgagacagtggaatgtttacacatcccaaaagcttggaatttggcgggtggaatgc 3276  
QY 3688 cttaattgagagatctacatgcttggatgagacacttttcttggtgtacgaacaaataag 3747  
Db 3277 cttaattgagagatctacatgcttggatgagacacttttcttggtgtacgaacaaataag 3336  
QY 3748 ttgtcagaatcccttgggaatggtcttgcacacattagagaatttcaagatcacagcttca 3807  
Db 3337 ttgtcagaatcccttgggaatggtcttgcacacattagagaatttcaagatcacagcttca 3396  
QY 3808 ggaacaatgagacagtggtggcccaaaagcttgcagagcttatattccggaataatcaat 3867  
Db 3397 ggaacaatgagacagtggtggcccaaaagcttgcagagcttatattccggaataatcaat 3456  
QY 3868 gctctgagacacccaagagagccctttcttgatcaaggttgatctgttgcaccaatgat 3927  
Db 3457 gctctgagacacccaagagagccctttcttgatcaaggttgatctgttgcaccaatgat 3516  
QY 3928 attcaagacatcaagaacagaggtgcctgcagaaagtctccagcctctacatctcag 3987  
Db 3517 attcaagacatcaagaacagaggtgcctgcagaaagtctccagcctctacatctcag 3576  
QY 3988 ttatatacatgatagtcttcttgatggaaagaatgtagagacttatccggaataatccact 4047  
Db 3577 ttatatacatgatagtcttcttgatggaaagaatgtagagacttatccggaataatccact 3636  
QY 4048 ggaaccttaatggtctctcttcttgacaatgtagatcatccttgataaacaataatttt 4107  
Db 3637 ggaaccttaatggtctctcttcttgacaatgtagatcatccttgagataaacaataatttt 3696  
QY 4108 aaacctccaattatgtctcogatacatccogttgcacccaactcaattatagcatctcagc 4167  
Db 3697 aaacctccaattatgtctcogatacatccogttgcacccaactcaattatagcatctcagc 3756  
QY 4168 actcttcgcagtggaagtgtgagctgttgaatttaaatagtgcagagcttcgcatgggaatg 4227  
Db 3757 actcttcgcagtggaagtgtgagctgttgaatttaaatagtgcagagcttcgcatgggaatg 3816  
QY 4228 gagaatgaagcaatatcaagatgacagatatacgtcttcaacacacttcaacataatgatt 4287  
Db 3817 gagaatgaagcaatatcaagatgacagatatacgtcttcaacacacttcaacataatgatt 3876  
QY 4288 gccacctgtctctcttcaaaaagctgcagctcacctccaagggaggttaatgcctggaga 4347  
Db 3877 gccacctgtctctcttcaaaaagctgcagctcacctccaagggaggttaatgcctggaga 3936  
QY 4348 cctcaggtgaataatcaaaaaggttgctggcaagtggacttccaagaagaatgaaatgc 4407  
Db 3937 cctcaggtgaataatcaaaaaggttgctggcaagtggacttccaagaagaatgaaatgc 3996  
QY 4408 acagagtaactactcaagggagtaaaatcctctgcttaccagacatgtaatgaaagattc 4467  
Db 3997 acagagtaactactcaagggagtaaaatcctctgcttaccagacatgtaatgaaagattc 4056  
QY 4468 ctcaatctcaagcagtcagaatgagcaltcagtggaactccttcttcaagaatgcaaaagta 4527  
Db 4057 ctcaatctcaagcagtcagaatgagcaltcagtggaactccttcttcaagaatgcaaaagta 4116  
QY 4528 aaggttttccagggaaataaagaatccttccacacgtgttggaactcttcaagccacgg 4587  
Db 4117 aaggttttccagggaaataaagaatccttccacacgtgttggaactcttcaagccacgg 4176  
QY 4588 ttactagctcgtacacttgcgaattccaccccaaggttgggtgacccaagattgcctcagag 4647  
Db 4177 ttactagctcgtacacttgcgaattccaccccaaggttgggtgacccaagattgcctcagag 4236  
QY 4648 atggaggtctcgtggtcgtgagacagaacactctac 4683  
Db 4237 atggaggtctcgtggtcgtgagacagaacactctac 4272



RESULT 14

AAx82258  
ID AAX82258 standard; cDNA: 4373 BP.

XX AAX82258;

XX 18-AUG-1999 (first entry)

DE Beta-domain deleted Factor VIII protein encoding cDNA.

XX Factor VIII protein; gene modification; gene therapy; clinical disorder;  
KW splicing pattern; RNA processing; gene regulation; beta-domain; human;  
KM ss.

XX Homo sapiens.

PN W09929848-A1.

XX 17-JUN-1999.

PF 25-NOV-1998: 98MO-US5354.

PR 16-JAN-1998: 98US-0071596.

PR 05-DEC-1997: 97US-0067614.

PA (IMMU-) IMMUNE RESPONSE CORP.

PI Bldingmaier S, Gonzales JEN, Ill CR, Yang CQ;

XX WPI: 1999-385602/32.

DR P-PSDB; AAY21675.

PT Genes and vectors exhibiting increased expression and novel splicing  
PT patterns, useful for expression of, e.g. beta-domain deleted factor  
PT VIII

PS Claim 9; Page 72-78; 123pp; English.

CC The invention describes novel genes and vectors exhibiting increased  
CC expression and novel splicing patterns. It provides a gene encoding a  
CC Factor VIII protein, that comprises one or more consensus or near  
CC consensus splice sites which have been corrected to increase expression.  
CC The method, DNA sequences and expression vectors can be used to increase  
CC the expression of a gene, especially a Factor VIII gene. Genes containing  
CC modified 5' and/or 3' untranslated regions have optimized expression  
CC levels and tissue-specific expression. The methods are used for  
CC identification and correction of consensus splice sites, addition of  
CC introns, optimization of 5' and 3' untranslated regions and increase in  
CC cytoplasmic RNA accumulation. Hence the DNAs are useful in gene therapy  
CC to treat a clinical disorder, to study RNA processing and/or gene  
CC regulation. The present sequence represents a cDNA encoding a beta-domain  
CC deleted factor VIII protein.

XX Sequence 4373 BP; 1278 A; 1011 C; 938 G; 1146 T; 0 other;

Query Match 87.3%; Score 4161; DB 20; Length 4373;  
Best Local Similarity 96.8%; Pred. No. 0;  
Matches 42/3; Conservative 0; Mismatches 100; Indels 42; Gaps 1;

QY 271 atgcaaatagagctcaccctgtctctctgtcctttgtcctttgtcagatctgtcttagtgc 330

DB 1 atggaatagagctcaccctgtctctctgtcctttgtcctttgtcagatctgtcttagtgc 60

QY 331 accagaagatactactcctgggtgcagtggaactgtcatggaactatgcaagaatgatctc 390

DB 61 accagaagatactactcctgggtgcagtggaactgtcatggaactatgcaagaatgatctc 120

QY 391 ggtgagctgctgtgagcgaagaattcctcctcctagagtgccaatactcttccattcaac 450

DB 121 ggaagagctgctgtgagcgaagaattcctcctcctcgtcgtgccaatactcttccattcaac 180

QY 451 acctcagtcgtgtacaaaaaagactctgtttgttagaatccaggttcaacctttcaacatc 510

DB 181 acctcagtcgtgtacaaaaaagactctgtttgttagaatccaggttcaacctttcaacatc 240

QY 511 gctaaagcgaagccacctgtgagtgtctgtctaggtctccatccagctgaggtttat 570

DB 241 gctaaagcgaagccacctgtgagtgtctgtctaggtctccatccagctgaggtttat 300

QY 571 gatacagtggtcattacacttaagaacatgctcccatccttcaagcttcaagcttcaagct 630

DB 301 gatacagtggtcattacacttaagaacatgctcccatccttcaagcttcaagcttcaagct 360

QY 631 ggttattcctacttgaaagactctctgaggagcttgaaatagtatgatacagacagtc 690

DB 361 ggttattcctacttgaaagactctctgaggagcttgaaatagtatgatacagacagtc 420

QY 691 gagaagaagaatgataagctcttccctgtgtggaagccatcatatgtctgtgcaggtccg 750

DB 421 gagaagaagaatgataagctcttccctgtgtggaagccatcatatgtctgtgcaggtccg 480

QY 751 aagaagaatgtgccaatgtgcctctgacccaactgtgcttcaactcatatcttctcat 810

DB 481 aagaagaatgtgccaatgtgcctctgacccaactgtgcttcaactcatatcttctcat 540

QY 811 gtgacctgtgtataaagacttgaaattcaggccctcatgtgagccctactagatgtagaaa 870

DB 541 gtgacctgtgtataaagacttgaaattcaggccctcatgtgagccctactagatgtagaaa 600

QY 871 gggagctctggcccaaggaaagacacagaactcttgacaaaatttactactctttgtctga 930

DB 601 gggagctctggcccaaggaaagacacagaactcttgacaaaatttactactctttgtctga 660

QY 931 ttgatgaaggaaagttgtgcactcagaaacaagaaccccttgatgacagatgagat 990

DB 661 ttgatgaaggaaagttgtgcactcagaaacaagaaccccttgatgacagatgagat 720

QY 991 gctgcattcgtctcggtcgtggtcctaaatgacacagtcgaatgtatgtataaagaagctc 1050

DB 721 gctgcattcgtctcggtcgtggtcctaaatgacacagtcgaatgtatgtataaagaagctc 780

QY 1051 ctgcagagctcgtatgtatgatacgaagaataatcagtcattgtgcagatgtatgaaatgggc 1110

DB 781 ctgcagagctcgtatgtatgatacgaagaataatcagtcattgtgcagatgtatgaaatgggc 840

QY 1111 acctcctggaagagcactcaataatctcctgaaagtgcaacacattctctgtgaaagacct 1170

DB 841 acctcctggaagagcactcaataatctcctgaaagtgcaacacattctctgtgaaagacct 900

QY 1171 cgcagagctccttggaataatctgcgaataaacttctctactcgtcgaacacactctgtatg 1230

DB 901 cgcagagctccttggaataatctgcgaataaacttctctactcgtcgaacacactctgtatg 960

QY 1231 gaccttgacaagtttctacgtgtttgtcatatccttcccaacaagtgtgcatgaaa 1290

DB 961 gaccttgacaagtttctacgtgtttgtcatatccttcccaacaagtgtgcatgaaa 1020

QY 1291 gcttatgtcaagaatgacagctgtccagaggaaccccaactacgaatgataaataatgaa 1350

DB 1021 gcttatgtcaagaatgacagctgtccagaggaaccccaactacgaatgataaataatgaa 1080

QY 1351 gaagcgaagaactatgtatgatacttacttgatctgaaatgtgagtgtgcaggtttgat 1410

DB 1081 gaagcgaagaactatgtatgatacttacttgatctgaaatgtgagtgtgtgcagatttgat 1140

QY 1411 gatgaacaactcctccttccatccaaatcgcctcagttgtccaaagaagatccttaaat 1470

DB 1141 gatgaacaactcctccttccatccaaatcgcctcagttgtccaaagaagatccttaaat 1200

QY 1471 tgggtacattacattgtctcgtcgaaggaaggaactcgggaagatactccttaagccctgcgc 1530

DB 1201 tgggtacattacattgtctcgtcgaaggaaggaactcgggaagatactccttaagccctgcgc 1260

QY 1531 ccgcatgacagaagtataaaagtcaatatttgaacaatgtgccctcagcgattgtgtaag 1590

Db 1261 cccgatcacgaagttaataaaglcataatttgacaatgcccctcagcgattggaag 1320  
 QY 1591 aagracaaaaaagtcgcatattatgcatcacacgatagaaacctttaagactcgtgaagc 1650  
 Db 1321 aagtacaaaaaagtcgcatattatgcatcacacgatagaaacctttaagactcgtgaagc 1380  
 QY 1651 attacagatgaatcagaagatcttggaccttctacttatggggaagtggagacacatg 1710  
 Db 1381 attacagatgaatcagaagatcttggaccttctacttatggggaagtggagacacatg 1440  
 QY 1711 ttgattatattaagaaatcagaacgacacatataacatctacccctcaggaatcact 1770  
 Db 1441 ctcatattattaagaaatcagaacgacacatataacatctacccctcaggaatcacc 1500  
 QY 1771 gatgctcgctctcttcttcttaagaagagatccaaaaggtgtaaaccttgaagatctt 1830  
 Db 1501 gatgctcgctctcttcttcttaagaagagatccaaaaggtgtaaaccttgaagatctt 1560  
 QY 1831 ccaattctgcaggaagaatattccaataataaattgacagtgactgtagaagatgggcca 1890  
 Db 1561 ccaattctgcaggaagaatattccaataataaattgacagtgactgtagaagatgggcca 1620  
 QY 1891 actaaatcagatcctcgtgctgcctgacccgctattactctagttctgccaatgtgagaga 1950  
 Db 1621 actaaatcagatcctcgtgctgcctgacccgctattactctagttctgccaatgtgagaga 1680  
 QY 1951 gatcagatctcagagatcatttggcctctccatctcgtacaaaagaatcgttagatcaa 2010  
 Db 1681 gatcagatctcagagatcatttggcctctccatctcgtacaaaagaatcgttagatcaa 1740  
 QY 2011 agggaaacacagataatgttcagacaagaagaatgctatcctgttcttctgatttgagag 2070  
 Db 1741 agggaaacacagataatgttcagacaagaagaatgctatcctgttcttctgatttgagag 1800  
 QY 2071 aacogaaagctgtgtaacctcacagaaatatacaagcttctccccaatccagcttggagtg 2130  
 Db 1801 aacogaaagctgtgtaacctcacagaaatatacaagcttctccccaatccagcttggagtg 1860  
 QY 2131 cagcttgagagatccagagatcccaagctccacaatcatatgacagcatcaatgtgtatgtt 2190  
 Db 1861 cagcttgagagatccagagatcccaagctccacaatcatatgacagcatcaatgtgtatgtt 1920  
 QY 2191 ttgagatggttgagatgctgtcagcttcttgcacatgagtgagacacatgcatcttaagc 2250  
 Db 1921 ttgagatggttgagatgctgtcagcttcttgcacatgagtgagacacatgcatcttaagc 1980  
 QY 2251 attggagacagacatgactcctctctgtcttctctctcgtgatatataccttcaacaacaa 2310  
 Db 1981 attggagacagacatgactcctctctgtcttctctctcgtgatatataccttcaacaacaa 2040  
 QY 2311 atggtctatagaagaacacacaccccatattccatctcagagagaacatgtcttcatgtctg 2370  
 Db 2041 atggtctatagaagaacacacaccccatattccatctcagagagaacatgtcttcatgtctg 2100  
 QY 2371 atggaacacccgaagctatgagatcttggggtgagcacaaacccagacttctggaaacagagc 2430  
 Db 2101 atggaacacccgaagctatgagatcttggggtgagcacaaacccagacttctggaaacagagc 2160  
 QY 2431 atgacgaccttactgaaggttctcagatattgtgacaagaacacatgctgattatatacagagac 2490  
 Db 2161 atgacgaccttactgaaggttctcagatattgtgacaagaacacatgctgattatatacagagac 2220  
 QY 2491 agttatagaagatatttcagatattctgctgagtaaaaaacatgacatctgaacaaagaagc 2550  
 Db 2221 agttatagaagatatttcagatattctgctgagtaaaaaacatgacatctgaacaaagaagc 2280  
 QY 2551 ttctccagaaatccaagacacccctagacatagcaaaagaatcttaatgacccccacca 2610  
 Db 2281 ttctccagaa-----ccccacca 2298  
 QY 2611 gtcttgaacgcacatcaacgggaataaactcgtactacttctcagatcagaatcaagaagaa 2670  
 |||||||

Db 2299 gtcttgaacgcacatcaacgggaataaactcgtactactcttccatcatcagaatcaagaagaa 2358  
 QY 2671 attgacatgatgatattccatatcatagtttgaataagaagaagaatttgcatttatgat 2730  
 Db 2359 attgacatgatgatattccatatcatagtttgaataagaagaagaatttgcatttatgat 2418  
 QY 2731 gaagatgaataatcagaagcccccgagcttccaaaagaagaacacagacatctttatgtct 2790  
 Db 2419 gaagatgaataatcagaagcccccgagcttccaaaagaagaacacagacatctttatgtct 2478  
 QY 2791 gcaagtgaagaagctcgtggaattatggatgagtagctccccaatgltctaaagaacag 2850  
 Db 2479 gcaagtgaagaagctcgtggaattatggatgagtagctccccaatgltctaaagaacag 2538  
 QY 2851 gctcagaagtgcagtgctccctcagttccagaagaatggttctccagaatcttatgtgct 2910  
 Db 2539 gctcagaagtgcagtgctccctcagttccagaagaatggttctccagaatcttatgtgct 2598  
 QY 2911 tcccttactcagcccttatataccgttggagaactaaatgaaacatttgggactccctgggcca 2970  
 Db 2599 tcccttactcagcccttatataccgttggagaactaaatgaaacatttgggactccctgggcca 2658  
 QY 2971 tatataagaacagaagttgaaagataatcatatgtagtaacttccagaaatcagagcctcgt 3030  
 Db 2659 tatataagaacagaagttgaaagataatcatatgtagtaacttccagaaatcagagcctcgt 2718  
 QY 3031 cccctattccttctatctctagcttattcttataggaagaatcagaggaacagaa 3090  
 Db 2719 cccctattccttctatctctctccatatacataggaagaatcagaggaacagaa 2778  
 QY 3091 cctagaanaaacttctcagagcctaatgaaacaaacacttcttggaaagtgcacacat 3150  
 Db 2779 cctagaanaaacttctcagagcctaatgaaacaaacacttcttggaaagtgcacacat 2838  
 QY 3151 cataatgcaacccataaagaatgagttgagctgcagaaagcctgagctattctctatgtt 3210  
 Db 2839 cataatgcaacccataaagaatgagttgagctgcagaaagcctgagctattctctatgtt 2898  
 QY 3211 gacctggaaaagaatgtgcactcagagcctgattggaaccccttctgtctgcacacacac 3270  
 Db 2899 gacctggaaaagaatgtgcactcagagcctgattggaaccccttctgtctgcacacacac 2958  
 QY 3271 acactgaacccctgcatatggagagaagaatgacatgacagaatcttctcgttttccac 3330  
 Db 2959 acactgaacccctgcatatggagagaagaatgacatgacagaatcttctcgttttccac 3018  
 QY 3331 atcttgaatgagaccaaagaagctgtgtaacttcaactgaataatggaagaacatgcaagct 3390  
 Db 3019 atcttgaatgagaccaaagaagctgtgtaacttcaactgaataatggaagaacatgcaagct 3078  
 QY 3391 cccctgcaatatccagatggaagaatcccaactttaaagaagaatatacgtctccatgaatc 3450  
 Db 3079 cccctgcaatatccagatggaagaatcccaactttaaagaagaatatacgtctccatgaatc 3138  
 QY 3451 aatggctacataatgataacatacacttggcttgaatgtgctcagaatcaaaagaatctga 3510  
 Db 3139 aatggctacataatgataacatacacttggcttgaatgtgctcagaatcaaaagaatctga 3198  
 QY 3511 tggatctgctcagcaatgggcaagcaatgaaacaatcatatctatcttcaatgagacat 3570  
 Db 3199 tggatctgctcagcaatgggcaagcaatgaaacaatcatatcttcaatctcccgagacat 3258  
 QY 3571 gttgttacatgtaagaaaaaagaagagataataaattgacatgtacaaatctctatccaggt 3630  
 Db 3259 gttgttacatgtaagaaaaaagaagagataataaattgacatgtacaaatctctatcccgga 3318  
 QY 3631 gttttgagaagaatggaagaatgtaacatccaagaagctggagaatttgggggtgagatgcctt 3690  
 Db 3319 gttttgagaagaatggaagaatgtaacatccaagaagctggagaatttgggggtgagatgcctt 3378  
 QY 3691 attggcgaagcatctacatgctggagatgagacacatttctcgtgtgtagacgacaataagtg 3750  
 Db 3379 attggcgaagcatctacatgctggagatgagacacatttctcgtgtgtagacccaataagtg 3438

```
OY 3751 cagatccctcgtggaatggtcttcgacacattagagatttcagattacagcttcagga 3810
    |||||||
DB 3439 cagatccctcgtggaatggtcttcgacacattagagatttcagattacagcttcagga 3498
OY 3811 caatatgacagctgggcccccaagctggccagactctctattccggatcacaatgcc 3870
    |||||||
DB 3499 caatatgacagctgggcccccaagctggccagactctctattccggatcacaatgcc 3558
OY 3871 tggagcccaaggagcccttcttcgatacgaagtggatccgttggcccaatgattat 3930
    |||||||
DB 3559 tggagcccaaggagcccttcttcgatacgaagtggatccgttggcccaatgattat 3618
OY 3931 caaggacatcaagaccaggggtgcgcgtcgaagaattctccagctctacatctcaatt 3990
    |||||||
DB 3619 caaggacatcaagaccaggggtgcgcgtcgaagaattctccagctctacatctcaatt 3678
OY 3991 atatcatgatatagcttcgtgagaagaagtgagcagacttatcgaaggaattccactgga 4050
    |||||||
DB 3679 atcatctgtatagcttcgtgagaagaagtgagcagacttatcgaaggaattccactgga 3738
OY 4051 accttaatggtctctcttcttgcaatgtgattcatctgggataaacaacatattttac 4110
    |||
DB 3739 acctcatggtctctcttcttgcaatgtgattcatctgggataaacaacatattttac 3798
OY 4111 cctccaattatgtctgatatatcatcctgtttgacccaactcatatagatctgcagacct 4170
    |||
DB 3799 cctccaattatgtctgatatatcatcctgtttgacccaactcatatagatctgcagacct 3858
OY 4171 ctccgcagtagagtgatgggggtgtgatttaattagttgcaacagatcttgggaatggag 4230
    |||||||
DB 3859 ctccgcagtagagtgatgggggtgtgatttaattagttgcaacagatcttgggaatggag 3918
OY 4231 agtaaaagcaatcatcagatctgcacagattactctcatcctcctcttaccataatgtttcc 4290
    |||||||
DB 3919 agtaaaagcaatcatcagatctgcacagattactctcatcctcctcttaccataatgtttcc 3978
OY 4291 acctgtctccttccaagaagctgcagctcgaacctccaaggaggaatgatgtcctggagacct 4350
    |||||||
DB 3979 acctgtctccttccaagaagctgcagctcgaacctccaaggaggaatgatgtcctggagacct 4038
OY 4351 caggtgataatccaagaagtggtctcaagtgagcttccaggaagacatgaagaatgcaca 4410
    |||
DB 4039 caggtgataatccaagaagtggtctcaagtgagcttccaggaagacatgaagaatgcaca 4098
OY 4411 gggagtaactactcagagtgaaatactctgtcttacagcagatgatagtgaaggaattcttc 4470
    |||
DB 4099 gggagtaactactcagagtgaaatactctgtcttacagcagatgatagtgaaggaattcttc 4158
OY 4471 atctccagcagtcagaatgagcagatcagtgagctctcttttttcagaatgagcaagaatgaag 4530
    |||
DB 4159 atatcgtctcgtcgaagaatgagcagatcagtgagctctcttttttcagaatgagcaagaatgaag 4218
OY 4531 gtttttcagggaatacgaagctccttcacagcttggatgaaatctcttagaccacccggtta 4590
    |||
DB 4219 gtttttcagggaatacgaagctccttcacagcttggatgaaatctcttagaccacccggtta 4278
OY 4591 ctgactcgtactactcgaattcaccaccccaagatgtgggtgacccagatgtcccttaagatg 4650
    |||
DB 4279 ctgactcgtactactcgaattcaccaccccaagatgtgggtgacccagatgtcccttaagatg 4338
OY 4651 gaggtctcgtggtcgtgagagcagacagacctactg 4685
    |||
DB 4339 gaggtctcgtggtcgtgagagcagacagacctactg 4373
```

## RESULT 15

AAN81544

ID AAN81544 standard; DNA: 4830 BP.

XX AAN81544;  
AC  
XX  
DT 04-DEC-1990 (first entry)

```
XX XX
DE Human Factor VIII-C analog having exon 14 deleted.
XX
KW Human Factor VIII-C analog; exon deletion; coagulation disorders;
XX haemophilia; ss.
XX
PN EP265778-A.
XX
PD 04-MAY-1988.
XX
PE 14-OCT-1987; 87EP-0115043.
XX
PR 15-OCT-1986; 86US-0919153.
XX
PA (RORE ) RORER INT OVERSEAS.
XX
PI Sarver N, Drohan W.
XX
DR WPI; 1988-120930/18.
XX
PT Human Factor VIII-C analogue free of other proteins - produced from
PS recombinant DNA and used in the treatment of coagulation disorders.
XX
PS Claim 3; Page 13; 42pp; English.
XX
CC The genetically engineered analogue can provide a dependable and
CC readily available therapeutic agent to be used in the treatment of
CC haemophilia and coagulation disorders in humans or animals.
CC See also AAN81543-45.
XX
SQ Sequence 4830 BP; 1392 A; 1091 C; 1058 G; 1289 T; 0 other;

Query Match 87.1%; Score 4152.2; DB 9; Length 4830;
Best Local Similarity 94.8%; Pred. No. 0;
Matches 4386; Conservative 0; Mismatches 43; Indels 196; Gaps 1;

OY 271 atgcaaatgagacgtctcacactgtctcttctctgtgcctttggagatctgattagtgcc 330
    |||||||
DB 1 atgcaaatgagacgtctcacactgtctcttctctgtgcctttggagatctgattagtgcc 60
OY 331 accggaagatctactcctcgtgtgacagtggaactgtcagatgcaggaactatagcaaatgatctcc 390
    |||||||
DB 61 accggaagatctactcctcgtgtgacagtggaactgtcagatgcaggaactatagcaaatgatctcc 120
OY 391 ggtgagctgcctgtggaacgcaagaattccctcctagagtgccaatccttttcattcaaac 450
    |||||||
DB 121 ggtgagctgcctgtggaacgcaagaattccctcctagagtgccaatccttttcattcaaac 180
OY 451 acctcagtcgtgtacaaaagactcgtttgttagaattcaaggttcaaccttttcaaacatc 510
    |||||||
DB 181 acctcagtcgtgtacaaaagactcgtttgttagaattcaaggttcaaccttttcaaacatc 240
OY 511 gctaaagccaaggccacctgtgagtgctgtctagtgcttccatccacagctgaggtttat 570
    |||||||
DB 241 gctaaagccaaggccacctgtgagtgctgtctagtgcttccatccacagctgaggtttat 300
OY 571 gatacagtggtcatctaaccttaagaacatggtctccatcactgtcaggttcaatgtgtt 630
    |||||||
DB 301 gatacagtggtcatctaaccttaagaacatggtctccatcactgtcaggttcaatgtgtt 360
OY 631 ggtgtactactcgtgaaagctctcgaaggagctgaaatgatgatcatcagaaccagtcacaag 690
    |||||||
DB 361 ggtgtactactcgtgaaagctctcgaaggagctgaaatgatgatcatcagaaccagtcacaag 420
OY 691 gagaaagaagatgatataagctctccctggtgtggaagcacaatacatatgtctggaggtccctg 750
    |||||||
DB 421 gagaaagaagatgatataagctctccctggtgtggaagcacaatacatatgtctggaggtccctg 480
OY 751 aaagagaatggtccaatgtgccttgaccacactgtgccttaaccatccatcattcttcat 810
    |||||||
DB 481 aaagagaatggtccaatgtgccttgaccacactgtgccttaaccatccatcattcttcat 540
```

QY 811 gtggacctggtlaaagaacttgaaatlcagagcccatlbgagccctactagatgtagaagaa 870  
 |||||  
 Db 541 gtggacctggtlaaagaacttgaaatlcagagcccatlbgagccctactagatgtagaagaa 600  
 QY 871 gggagctctggccaaggaagaacagacagaccttgacaaattatactacttttgcgtga 930  
 |||||  
 Db 601 gggagctctggccaaggaagaacagacagaccttgacaaattatactacttttgcgtga 660  
 QY 931 ttgatgaagggaanaaagtltggacctcagaaacaaagaccttgatggcggataggat 990  
 |||||  
 Db 661 ttgatgaagggaanaaagtltggacctcagaaacaaagaccttgatggcggataggat 720  
 QY 991 gctgcactctctcgggccttgagccctaaatgcacacagtcacatgtgtatgtgaacaagttct 1050  
 |||||  
 Db 721 gctgcactctctcgggccttgagccctaaatgcacacagtcacatgtgtatgtgaacaagttct 780  
 QY 1051 ctgcacagctctgatatgtgagtcacaggaanaatcagtcattgtgcagtgtatgtgaatggc 1110  
 |||||  
 Db 781 ctgcacagctctgatatgtgagtcacaggaanaatcagtcattgtgcagtgtatgtgaatggc 840  
 QY 1111 acaactctgaagtgcacataatctcctgaaggtgcacacatttctgtgaggaagacat 1170  
 |||||  
 Db 841 acaactctgaaggtgcacataatctcctgaaggtgcacacatttctgtgaggaagacat 900  
 QY 1171 cgcgaagcgtctctggaatactgcacataacttcccttactgtctcaaacactctgagt 1230  
 |||||  
 Db 901 cgcgaagcgtctctggaatactgcacataacttcccttactgtctcaaacactctgagt 960  
 QY 1231 gaacctgagacagtttactactgttttgcatactctccacaaacatlgatgcatggaa 1290  
 |||||  
 Db 961 gaacctgagacagtttactactgttttgcatactctccacaaacatlgatgcatggaa 1020  
 QY 1291 gctatgtcaaaatgaagaacgtlccagagaagcccaactacagaaatgaanaaatatgaa 1350  
 |||||  
 Db 1021 gctatgtcaaaatgaagaacgtlccagagaagcccaactacagaaatgaanaaatatgaa 1080  
 QY 1351 gaagcgggaagacatgatgtatgtactactgatactgaaatgtgagtcagagttgcat 1410  
 |||||  
 Db 1081 gaagcgggaagacatgatgtatgtactactgatactgaaatgtgagtcagagttgcat 1140  
 QY 1411 gatgacaactctctcccttataccaaatctcgctcagttgccaagaagaaccttaaaact 1470  
 |||||  
 Db 1141 gatgacaactctctcccttataccaaatctcgctcagttgccaagaagaaccttaaaact 1200  
 QY 1471 ttgggtatactatcatgtgctgtgaaaggaagacttggagacttcccttaagccctcgcc 1530  
 |||||  
 Db 1201 ttgggtatactatcatgtgctgtgaaaggaagacttggagacttcccttaagccctcgcc 1260  
 QY 1531 cccgattgacagaagtataaaagtcacatattgaacaatlgagccctcagcgattgttag 1590  
 |||||  
 Db 1261 cccgattgacagaagtataaaagtcacatattgaacaatlgagccctcagcgattgttag 1320  
 QY 1591 aagttacaaaagtcgcgattatgycacatagatgaaacctttaagactcgttgaagct 1650  
 |||||  
 Db 1321 aagttacaaaagtcgcgattatgycacatagatgaaacctttaagactcgttgaagct 1380  
 QY 1651 attcagagatgaatcaggaactctggagacctttagtattgaggaagtltgagagacacgt 1710  
 |||||  
 Db 1381 attcagagatgaatcaggaactctggagacctttagtattgaggaagtltgagagacacgt 1440  
 QY 1711 ttgatataatttaagaatcagaagcagacacatataacatccctcagcgatcaact 1770  
 |||||  
 Db 1441 ttgatataatttaagaatcagaagcagacacatataacatccctcagcgatcaact 1500  
 QY 1771 gatgtccgtctcttgcattcacaaggagattacacaaaggtgtlaaacaatttgaagatttt 1830  
 |||||  
 Db 1501 gatgtccgtctcttgcattcacaaggagattacacaaaggtgtlaaacaatttgaagatttt 1560  
 QY 1831 ccaattctgcaggaagaataatcacaataataaattgaacagtgagctgtagaagaatggcca 1890  
 |||||  
 Db 1561 ccaattctgcaggaagaataatcacaataataaattgaacagtgagctgtagaagaatggcca 1620  
 QY 1891 actaaatcagatcctctcggtgcctgaccggtatctactctgattcgttlaatatgagaga 1950

Db 1621 actaaatcagatcctctcggtgcctgagccggtatctactctgattcgttlaatatgagaga 1680  
 |||||  
 QY 1951 gatctagcttcaagactcatcttgagccctctccatctgtctcaaaagaatctgtatagcaa 2010  
 |||||  
 Db 1681 gatctagcttcaagactcatcttgagccctctccatctgtctcaaaagaatctgtatagcaa 1740  
 QY 2011 agaggaacacagataatgtatagaacaagagaatgtacatccctgttttctgtatgtagag 2070  
 |||||  
 Db 1741 agaggaacacagataatgtatagaacaagagaatgtacatccctgttttctgtatgtagag 1800  
 QY 2071 aacccgaagctgtgatacctcaagagaatatalacaagcttltccccaatccagctgtgagtg 2130  
 |||||  
 Db 1801 aacccgaagctgtgatacctcaagagaatatalacaagcttltccccaatccagctgtgagtg 1860  
 QY 2131 cagcttgaggaatccagagttccaaagctccacaatcatgtcagaagcatcaatgtgtatgtt 2190  
 |||||  
 Db 1861 cagcttgaggaatccagagttccaaagctccacaatcatgtcagaagcatcaatgtgtatgtt 1920  
 QY 2191 ttgatgaatttgaggttgcagtttgttggcatgaggttggcagactggtacatctgaagc 2250  
 |||||  
 Db 1921 ttgatgaatttgaggttgcagtttgttggcatgaggttggcagactggtacatctgaagc 1980  
 QY 2251 attgagacagagctgaactcctcttctgtcttctctctgtgatatataccttcaacaacaa 2310  
 |||||  
 Db 1981 attgagacagagctgaactcctcttctgtcttctctctgtgatatataccttcaacaacaa 2040  
 QY 2311 atvgctatagaagaacaacactcaacctatcccatlctcagagaacactgtctcatgtctg 2370  
 |||||  
 Db 2041 atvgctatagaagaacaacactcaacctatcccatlctcagagaacactgtctcatgtctg 2100  
 QY 2371 atggaanaacccaagttatgatgttctgggttggccaacactcaacttctggaacaagagc 2430  
 |||||  
 Db 2101 atggaanaacccaagttatgatgttctgggttggccaacactcaacttctggaacaagagc 2160  
 QY 2431 atgacgccttacttgaaggttltctagtttggacaagaacacttgtatatactcagagac 2490  
 |||||  
 Db 2161 atgacgccttacttgaaggttltctagtttggacaagaacacttgtatatactcagagac 2220  
 QY 2491 agttatagaagatatltccaagatactgtcgtgagtraaaaacaatgtccaatgaacaagaagc 2550  
 |||||  
 Db 2221 agttatagaagatatltccaagatactgtcgtgagtraaaaacaatgtccaatgaacaagaagc 2280  
 QY 2551 tttctccagaatctca----- 2665  
 |||||  
 Db 2281 tttctccagaatctcaatctatgtgactctctgtctgtggaataccaatagtactcagata 2340  
 |||||  
 QY 2566 ----- 2565  
 Db 2341 ccaaaagaagagtggaatccccaagaaggtcaaccagaanaaacagcttlttaagaanaag 2400  
 |||||  
 QY 2566 ----- 2565  
 Db 2401 gatcacatttgcctctgaagcgtltgaaagcaatcatgcaatlagcaacaataatgag 2460  
 |||||  
 QY 2566 ----- agcacacctgacgtgagcagaagcaa 2592  
 |||||  
 Db 2461 ggaacaaataagccgaaatagaagtcacactggtggccacgaaggttagagactgaaagcgtg 2520  
 |||||  
 QY 2593 tttaatgccaaccccaacagttcttgaagaacgcatcaacgggaataactcgttactactct 2652  
 |||||  
 Db 2521 tgtctccaaaacccacacagttcttgaagaacgcatcaacgggaataactcgttactactct 2580  
 QY 2653 cagtcagatccaagaggaatctgactatgtatgatacatatcagttcagttgaanaatgaagaagaa 2712  
 |||||  
 Db 2581 cagtcagatccaagaggaatctgactatgtatgatacatatcagttcagttgaanaatgaagaagaa 2640  
 QY 2713 gatttgaacatttatatagaatgtgaanaatcagagcccgagcgtttaaagaagaagaa 2772  
 |||||  
 Db 2641 gatcttgacattatagatgagatgaaatcagagcccgagcgtttaaagaagaagaa 2700  
 |||||  
 QY 2773 cgaacatttatactgtcagtgagagagccttgagattatgagatgagtagctccca 2832  
 |||||

Db 2701 cgacactatttatttctgtgagtgagagagctctggagttatbgygaatgagtagctcccca 2760  
QY 2833 catgtcttaagaaacagagctcagagtgagtgctccctcagttcaagaagtgttttc 2892  
Db 2761 catgtcttaagaaacagagctcagagtgagtgctccctcagttcaagaagtgttttc 2820  
QY 2893 caggaatttactgtatgctccttactcagccctataccgtbgygaacataaagaacat 2952  
Db 2821 caggaatttactgtatgctccttactcagccctataccgtbgygaacataaagaacat 2880  
QY 2953 ttggagctccggggccatatataagacagaagttagaataatataatgttaacttc 3012  
Db 2881 ttggagctccggggccatatataagacagaagttagaataatataatgttaacttc 2940  
QY 3013 agaatacaggcctctcgtcccttactccttacttactaagcctatttcttatbgygaagat 3072  
Db 2941 agaatacaggcctctcgtcccttactccttacttactaagcctatttcttatbgygaagat 3000  
QY 3073 cagagcagaagagcagaacctagaaaaaacttgcacagcctaatabgaacaaacttac 3132  
Db 3001 cagagcagaagagcagaacctagaaaaaacttgcacagcctaatabgaacaaacttac 3060  
QY 3133 ttcttggaagatgcaacatcatatataatgcaaccataaagaatgagtttgaactgcaaaagctcg 3192  
Db 3061 ttcttggaagatgcaacatcatatataatgcaaccataaagaatgagtttgaactgcaaaagctcg 3120  
QY 3193 gcttatttctctgaatgttgaaccttgaaaaaagtgtgcaactcagagcctgattggaacctt 3252  
Db 3121 gcttatttctctgaatgttgaaccttgaaaaaagtgtgcaactcagagcctgattggaacctt 3180  
QY 3253 ctgtgtctgcaacactaaacacacatgaaacctgtcatagggagacaagtgcagtaagaa 3312  
Db 3181 ctgtgtctgcaacactaaacacacatgaaacctgtcatagggagacaagtgcagtaagaa 3240  
QY 3313 ttgtgcttgtttttcacacatctttagatgagacaaagctgtgatttacttcaacgaaatgt 3372  
Db 3241 ttgtgcttgtttttcacacatctttagatgagacaaagctgtgatttacttcaacgaaatgt 3300  
QY 3373 gaaagaacatgcaagagctccctcgcaataataccagaatgaaagatacccaactttaagaagat 3432  
Db 3301 gaaagaacatgcaagagctccctcgcaataataccagaatgaaagatacccaactttaagaagat 3360  
QY 3433 tatcgcttccatgcaataatgctgtacataatagatatcacacttgcttgaatgct 3492  
Db 3361 tatcgcttccatgcaataatgctgtacataatagatatcacacttgcttgaatgct 3420  
QY 3493 caggaatcaagagatcgatgtatcgtcagacatgagcagcaatgaaataccatctt 3552  
Db 3421 caggaatcaagagatcgatgtatcgtcagacatgagcagcaatgaaataccatctt 3480  
QY 3553 attcatttcagtgagacatgtgttcaactgtacgaaaaaagaagagatataaattgacatg 3612  
Db 3481 attcatttcagtgagacatgtgttcaactgtacgaaaaaagaagagatataaattgacatg 3540  
QY 3613 tacaactctcatccagtggtttttgagaacagtggaatgtttacatccaagctggaaat 3672  
Db 3541 tacaactctcatccagtggtttttgagaacagtggaatgtttacatccaagctggaaat 3600  
QY 3673 tggcgggtggaatgaccttatgtgagacatctacatcgtggagtgagacactttttctg 3732  
Db 3601 tggcgggtggaatgaccttatgtgagacatctacatcgtggagtgagacactttttctg 3660  
QY 3733 gtgtacagcaataagtgatcagactcccttggaatggtcttcttgacacataatagatatt 3792  
Db 3661 gtgtacagcaataagtgatcagactcccttggaatggtcttcttgacacataatagatatt 3720  
QY 3793 cagattacagcttcaggaacatatgacagtgggcccacaaagctgcgcaagcttcatat 3852  
Db 3721 cagattacagcttcaggaacatatgacagtgggcccacaaagctgcgcaagcttcatat 3780  
QY 3853 tccggatcaatcaatgctgtagagaccaaagagccttttcttgatcaaatgtgagctg 3912  
Db 3781 tccggatcaatcaatgctgtagagaccaaagagccttttcttgatcaaatgtgagctg 3840

QY 3913 ttggccaatgatatttcaacgacatcaagaaccaggttgcccgatcagaagtctccagc 3972  
Db 3841 ttggccaatgatatttcaacgacatcaagaaccaggttgcccgatcagaagtctccagc 3900  
QY 3973 ctctacatctcagattatcatcatatgatatgcttgatbgygaagatgagcaacttat 4032  
Db 3901 ctctacatctcagattatcatcatatgatatgcttgatbgygaagatgagcaacttat 3960  
QY 4033 cgaaggaattccactgggaaccttaatggttcttcttggaatgagatgatatcgggata 4092  
Db 3961 cgaaggaattccactgggaaccttaatggttcttcttggaatgagatgatatcgggata 4020  
QY 4093 aaacacaaatattttaaccctccaattatgtctgataacatccgttgcacccaactcat 4152  
Db 4021 aaacacaaatattttaaccctccaattatgtctgataacatccgttgcacccaactcat 4080  
QY 4153 tatagacttcgacgacctcttcgcaatgagatgagatgagctgtgatattaaatagttgacg 4212  
Db 4081 tatagacttcgacgacctcttcgcaatgagatgagatgagctgtgatattaaatagttgacg 4140  
QY 4213 atgcgaattgggaatgggaagttaaagcaatatcagatgcacagattactgttcaatccctac 4272  
Db 4141 atgcgaattgggaatgggaagttaaagcaatatcagatgcacagattactgttcaatccctac 4200  
QY 4273 ttaccaaatagtttgccaacctgttctccttcaaaaagctcgacttcaaccccaagggag 4332  
Db 4201 ttaccaaatagtttgccaacctgttctccttcaaaaagctcgacttcaaccccaagggag 4260  
QY 4333 agtaatgcttgagacactcagatgtaataatccaaaagatgagctgcaagtgtaccag 4392  
Db 4261 agtaatgcttgagacactcagatgtaataatccaaaagatgagctgcaagtgtaccag 4320  
QY 4393 aagaacaatggaatgcacaggaagttaacttcaaggagtaaaatctcgttaccagaatg 4452  
Db 4321 aagaacaatggaatgcacaggaagttaacttcaaggagtaaaatctcgttaccagaatg 4380  
QY 4453 tatgtgaagaggttctcctatctccagcagatcaagaatgcaatcagtgactcttctttt 4512  
Db 4381 tatgtgaagaggttctcctatctccagcagatcaagaatgcaatcagtgactcttctttt 4440  
QY 4513 cagaatggaagaatgaaaggtttttcagaaggaaataaagactcctacacactgtggtgaac 4572  
Db 4441 cagaatggaagaatgaaaggtttttcagaaggaaataaagactcctacacactgtggtgaac 4500  
QY 4573 tctctagaagcccaacggttaactgactcgcttaacttgaattcaaccccaaggttgggtgac 4632  
Db 4501 tctctagaagcccaacggttaactgactcgcttaacttgaattcaaccccaaggttgggtgac 4560  
QY 4633 cagattgcccctgagatggaagtctctggtcgtgcaagcacaagacacttactagaggtg 4692  
Db 4561 cagattgcccctgagatggaagtctctggtcgtgcaagcacaagacacttactagaggtg 4620  
QY 4693 cgaagtctc 4699  
Db 4621 ccaactgctc 4627

Search completed: November 18, 2001, 05:25:05  
Job time: 37037 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2001, 01:27:15 ; Search time 22045.3 Seconds  
(Without alignments)  
3343.294 Million cell updates/sec

Title: US-09-689-430-1\_COPY\_150\_4914  
Perfect score: 4765  
Sequence: 1 cctcttcaagtaaacagta.....ttgggtcgttgcgcgac 4765

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_ba1: \*  
2: gb\_ba2: \*  
3: gb\_ba3: \*  
4: gb\_in1: \*  
5: gb\_in2: \*  
6: gb\_in3: \*  
7: gb\_om: \*  
8: gb\_ov: \*  
9: gb\_pat1: \*  
10: gb\_pat2: \*  
11: gb\_ph: \*  
12: gb\_pl1: \*  
13: gb\_pl2: \*  
14: gb\_pl3: \*  
15: gb\_pl4: \*  
16: em\_ba1: \*  
17: em\_ba2: \*  
18: em\_fun: \*  
19: em\_hgo\_hum: \*  
20: em\_hgo\_inv: \*  
21: em\_hgo\_rod: \*  
22: em\_htg\_hum1: \*  
23: em\_htg\_hum2: \*  
24: em\_htg\_hum3: \*  
25: em\_htg\_hum4: \*  
26: em\_htg\_hum5: \*  
27: em\_htg\_hum6: \*  
28: em\_htg\_hum7: \*  
29: em\_htg\_hum8: \*  
30: em\_htg\_inv1: \*  
31: em\_htg\_inv2: \*  
32: em\_htg\_other: \*  
33: em\_htg\_rod: \*  
34: em\_hum1: \*  
35: em\_hum2: \*  
36: em\_hum3: \*  
37: em\_hum4: \*  
38: em\_hum5: \*  
39: em\_hum6: \*  
40: em\_hum7: \*  
41: em\_in: \*  
42: em\_om: \*  
43: em\_or: \*

150-4835

44: em\_ov: \*  
45: em\_pat: \*  
46: em\_ph: \*  
47: em\_pl: \*  
48: em\_rod: \*  
49: em\_sts: \*  
50: em\_sy: \*  
51: em\_un: \*  
52: em\_vi: \*  
53: gb\_sts1: \*  
54: gb\_sts2: \*  
55: gb\_sts3: \*  
56: gb\_sy: \*  
57: gb\_un: \*  
58: gb\_v11: \*  
59: gb\_v12: \*  
60: gb\_htg1: \*  
61: gb\_htg2: \*  
62: gb\_htg3: \*  
63: gb\_htg4: \*  
64: gb\_htg5: \*  
65: gb\_htg6: \*  
66: gb\_htg7: \*  
67: gb\_htg8: \*  
68: gb\_htg9: \*  
69: gb\_htg10: \*  
70: gb\_htg11: \*  
71: gb\_htg12: \*  
72: gb\_htg13: \*  
73: gb\_htg14: \*  
74: gb\_htg15: \*  
75: gb\_htg16: \*  
76: gb\_htg17: \*  
77: gb\_htg18: \*  
78: gb\_htg19: \*  
79: gb\_htg20: \*  
80: gb\_htg21: \*  
81: gb\_htg22: \*  
82: gb\_htg23: \*  
83: gb\_htg24: \*  
84: gb\_htg25: \*  
85: gb\_pr1: \*  
86: gb\_pr2: \*  
87: gb\_pr3: \*  
88: gb\_pr4: \*  
89: gb\_pr5: \*  
90: gb\_pr6: \*  
91: gb\_pr7: \*  
92: gb\_pr8: \*  
93: gb\_pr9: \*  
94: gb\_rot1: \*  
95: gb\_rot2: \*  
96: gb\_in4: \*  
97: gb\_pr10: \*  
98: em\_ba3: \*

Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | DB | ID       | Description       |
|------------|--------|-------------|--------|----|----------|-------------------|
| 1          | 4411.6 | 92.6        | 9354   | 9  | AR003585 | AR003585 Sequence |
| 2          | 4357.6 | 91.5        | 4670   | 9  | AR110040 | AR110040 Sequence |
| 3          | 4148.8 | 87.1        | 4278   | 10 | I08644   | I08644 Sequence 4 |
| 4          | 4147.2 | 87.0        | 4281   | 10 | I08643   | I08643 Sequence 3 |
| 5          | 4085   | 85.7        | 4548   | 10 | I08642   | I08642 Sequence 2 |
| 6          | 4085   | 85.7        | 4551   | 10 | I08641   | I08641 Sequence 1 |
| 7          | 3748   | 78.7        | 5035   | 9  | AR034084 | AR034084 Sequence |
| 8          | 3748   | 78.7        | 5035   | 9  | AR071306 | AR071306 Sequence |

|    |        |      |       |    |            |
|----|--------|------|-------|----|------------|
| 9  | 3189.8 | 66.9 | 433.4 | 9  | AR029098   |
| 10 | 2605.4 | 54.7 | 445.1 | 9  | AR110041   |
| 11 | 2336.4 | 49.0 | 9009  | 9  | AR003710   |
| 12 | 2336.4 | 49.0 | 9009  | 9  | AR029065   |
| 13 | 2336.4 | 49.0 | 9009  | 10 | I31901     |
| 14 | 2336.4 | 49.0 | 9009  | 10 | 163424     |
| 15 | 2336.4 | 49.0 | 9009  | 97 | HNEV42II   |
| 16 | 2335.4 | 49.0 | 7056  | 10 | I27063     |
| 17 | 2334.8 | 49.0 | 8241  | 9  | A05328     |
| 18 | 2334.8 | 49.0 | 8241  | 9  | A07042     |
| 19 | 2334.8 | 49.0 | 8831  | 9  | E00527     |
| 20 | 2334.8 | 49.0 | 9029  | 97 | HNEV42II   |
| 21 | 2333.8 | 49.0 | 7272  | 10 | 105404     |
| 22 | 2333.8 | 49.0 | 8667  | 9  | AX052730   |
| 23 | 2333.2 | 49.0 | 8967  | 10 | I71409     |
| 24 | 2333.2 | 49.0 | 8967  | 93 | HSEV41IR   |
| 25 | 2333.2 | 49.0 | 8967  | 93 | HSEV41IR   |
| 26 | 2328.4 | 48.9 | 7440  | 10 | 108345     |
| 27 | 2314   | 48.6 | 7440  | 10 | 108457     |
| 28 | 2275.2 | 47.7 | 6999  | 10 | I71105     |
| 29 | 1848   | 38.8 | 7032  | 7  | AF016234   |
| 30 | 1802.8 | 37.8 | 7435  | 7  | AF049489   |
| 31 | 1802.8 | 37.8 | 7435  | 9  | AR003712   |
| 32 | 1802.8 | 37.8 | 7435  | 9  | AR029067   |
| 33 | 1802.8 | 37.8 | 7435  | 9  | 163427     |
| 34 | 1802.8 | 37.8 | 7435  | 9  | M05CVCV4II |
| 35 | 1752.2 | 36.8 | 6539  | 7  | SS049517   |
| 36 | 1751.2 | 36.8 | 6539  | 9  | AR029097   |
| 37 | 1651   | 33.6 | 1993  | 10 | I02047     |
| 38 | 1603.2 | 33.6 | 1993  | 10 | I02054     |
| 39 | 1240.4 | 26.0 | 1623  | 10 | I27064     |
| 40 | 991.4  | 20.8 | 3852  | 10 | I04400     |
| 41 | 990.8  | 20.8 | 3852  | 10 | E00422     |
| 42 | 893.4  | 18.7 | 1130  | 9  | AR003711   |
| 43 | 893.4  | 18.7 | 1130  | 9  | AR029066   |
| 44 | 893.4  | 18.7 | 1130  | 10 | I31900     |
| 45 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 46 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 47 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 48 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 49 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 50 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 51 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 52 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 53 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 54 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 55 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 56 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 57 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 58 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 59 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 60 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 61 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 62 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 63 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 64 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 65 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 66 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 67 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 68 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 69 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 70 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 71 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 72 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 73 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 74 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 75 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 76 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 77 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 78 | 893.4  | 18.7 | 1130  | 10 | 163426     |
| 79 | 893.4  | 1    |       |    |            |

## ALIGNMENTS

| RESULT     | 1   |               |             |        |  |             |
|------------|---|---------------|-------------|--------|--|-------------|
| LOCUS      | AR003585  |               |             |        |  |             |
| DEFINITION | Sequence  | 9354 bp       | DNA         |        |  | PAT         |
| ACCESSION  | AR003585  | 2 from patent | US 5744326. |        |  | 04-DEC-1998 |
| VERSION    | AR003585.1  | GI:3964844    |             |        |  |             |
| KEYWORDS   |   |               |             |        |  |             |
| SOURCE     | Unknown.  |               |             |        |  |             |
| ORGANISM   | Unknown.  |               |             |        |  |             |
| REFERENCE  | Unclassified.   |               |             |        |  |             |
| AUTHORS    | 1 (bases 1 to 9354)   |               |             |        |  |             |
| TITLE      | Ill,C.R. and Bidlingmeyer,S.                                      |               |             |        |  |             |
|            | use of viral cis-acting post-transcriptional regulatory sequences |               |             |        |  |             |
|            | to increase expression of intronless genes containing             |               |             |        |  |             |
|            | near-consensus splice sites                                       |               |             |        |  |             |
| JOURNAL    | Patent: US 5744326-A 2 28-APR-1998;                               |               |             |        |  |             |
| FEATURES   | Location/Qualifiers   |               |             |        |  |             |
| SOURCE     | 1..9354   |               |             |        |  |             |
|            | /organism="Unknown"   |               |             |        |  |             |
| BASE COUNT | 2506 a  | 2239 c        | 2161 g      | 2448 t |  |             |
| ORIGIN     |   |               |             |        |  |             |

|    |                       |   |               |                |              |         |
|----|-----------------------|---|---------------|----------------|--------------|---------|
|    | Query Match           | 92.6%   | Score 4411.6; | DB 9;          | Length 9334; |         |
|    | Best Local Similarity | 99.7%;  | Pred. No. 0;  | Mismatches 14; | Indels 0;    | Gaps 0. |
|    | Matches 4420;         | Conservative  | 0;            |                |              |         |
| Oy | 266                   | ccacatgcaaatagagcttcacactgcgtttctcgtgccttbtgcatltygatctgcta | 325           |                |              |         |
|    |                       |   |               |                |              |         |
|    |                       |   |               |                |              |         |
|    |                       |   |               |                |              |         |
|    |                       |   |               |                |              |         |
|    |                       |   |               |                |              |         |
| Db | 2960                  | CCACCATGTTTATAGAGCTTCACACTGGTCTTTCTGTGCGCTTTTGCAATTCTCGTTTA | 3019          |                |              |         |

|    |      |  |      |
|----|------|--|------|
| QY | 326  | gtgcaccagaagaatactactcgtggatgcagttggaacgtgcaatggagcctatgatcaaaatg    | 385  |
| Db | 3020 | GTGCCACCAAGAAATACTACCTGGGTGACATGGAAGACTGTCATGGGACTTTATGCAAAAGTG      | 3078 |
| QY | 386  | atcccgatgagctgacgctgtagcgcgaagaattccctcccttaagtagtccaaatccttccat     | 445  |
| Db | 3080 | ATTCGTGGTGAAGCTGCTGTGGAGCCAGAAATTTCTCTCTAGAGTGGCCAAATCTTTTCCAT       | 3139 |
| QY | 446  | tcaaacctcaagtcgtgtacaataaagaactgtttgttagaatcaacggttcacctttca         | 505  |
| Db | 3140 | TCAACACCTCTAGTGTGTACAAAAGACTGTGTTTGTAGAAATTCACAGGTGTCACTTTTCA        | 3199 |
| QY | 506  | acatcgcctaaagcccaagggccaccctggatgtgtctgttaagtgctccatcccaagcgttcagg   | 565  |
| Db | 3200 | ACATCGCTTAAGCCCAAGGCCACCTGGATGGGTGTGTAGGTCTTACCATCCAGGCTGGAG         | 3255 |
| QY | 566  | tctatgaataagtgatgcatataacttaagaacatggcttcccatccctgtcagttctcatg       | 625  |
| Db | 3260 | TTTATGTATACGTGTGTATTAACCTTAAGAACATGGCTTCCCATCTGTCACTCTTCAATG         | 3319 |
| QY | 626  | ctgtttgtagtaccctactgtgaaagctctctgaggagctgaatatgatatgtagcagcagtc      | 685  |
| Db | 3320 | CTGTGGTGTATCCTACTACTGGAAGACTCTGTGAGGAGCACTGAATATGATGTATCAGACAGTC     | 3379 |
| QY | 686  | aaagagggaagaagaaatgataaagaatctccctgtgtgaaagccataatattgtctggcagg      | 745  |
| Db | 3380 | AAAGGGAGAAAGAAATGATTAAGTCTTCCCTGTGTGAAGCCATATATATGTCTGGACGG          | 3439 |
| QY | 746  | tccctgaagaagaatgtagtccaaatgagcctctctaccacatgtgtccttaactatactctt      | 805  |
| Db | 3440 | TCTGTAAAGAGAAATGTGTCCAAATGGCCCTGTACCCACTGTAGCCTTACTCTCATATCTTT       | 3499 |
| QY | 806  | ctcaatgtgagacctgtgtataaagaacttgaattcaagcctcatatggagccctactagtatgta   | 865  |
| Db | 3500 | CTCATGTGTGACCTGTGTAAAAAGACTTTGATTTACAGGCTCATTTGGAGCCCTACATGATATGTA   | 3559 |
| QY | 866  | gagaaagggaatctctggccaagaagaacaaagaccccttgacaacattatatacacttttg       | 925  |
| Db | 3560 | GAGAAAGGAGTCTGGCCAGGAAGAAAGACACAGACCTTGGACACAAATTTATACTTTTGG         | 3619 |
| QY | 926  | ctgtattcttgataaggggaaagaagttggaactcagacaagaacaaagaactccctgtatgcagata | 985  |
| Db | 3620 | CTGTATTTTGAATGAAGGGAAAAAGTTGGGACATCAAAACAAAGAATCTGTGATGCAAGATA       | 3679 |
| QY | 986  | gggaatgctgcatctgtctcgggacctgtgacctaaatgcaacagtcacatggttatgtgaaca     | 1045 |
| Db | 3680 | GGGATGCTGCATCTGTCTCGGGCCTGTGGCTTAATATGCACACAGTCATGCTTATGTATAACA      | 3739 |
| QY | 1046 | ggtctctgcgcaggtctgatttggaatgccaagggaaatcgaatctatgtgagatgtagattgaa    | 1105 |
| Db | 3740 | GGTCTTGCCAGGCTGTGATTTGGATGTGCCACGGGAAATCAGTCTTATTTGGCAATGTATTTGGAA   | 3799 |
| QY | 1106 | tgaggaccacactccctgaagtgacatcaataatccctcgaaggtcacacattctctgtgaaga     | 1165 |
| Db | 3800 | TGGGACACACTCCGTGAAGTGCATCAATATATCTCGAAGTGCACACATTTCTTGTGAGGA         | 3855 |
| QY | 1166 | accatgcgcagcgctcctgtgaaatctcgcgaataacttcctactactgtcctaacaactct       | 1225 |
| Db | 3860 | ACCATGCCCGAGGCGCTCTTGTGAATATCAGCCATTAATCTTCCCTACTGCTCAAAACACTCT      | 3919 |
| QY | 1226 | tgaatgagcccttgagaagtttctctactgttttgtatataatctctctcccaagaatgtagtga    | 1285 |
| Db | 3920 | TGATGAGCACTTGGACACTTTTCTACTGTTTTGTATATTCCTTTCCACCAACATGTATGGCA       | 3979 |
| QY | 1286 | tgagaagcttatgtcaaaagtgaacagctgtctcgaaggaaaccccaactagaatgaaataa       | 1345 |
| Db | 3980 | TGGAAAGCTTATGTCTAAAGTGAAGAGCTGTCTCAGAGSAGAACCCCAACTACGAATGAAAAAFTA   | 4039 |
| QY | 1346 | atgaagaagcgggaagactatgatgatgatcttactgtatcttgaatgtagtgtgtcagagt       | 1405 |
| Db | 4040 | ATGAAGAAGCGGAAGACTATATGATGATGATCTTACTGATTTCTGAAATGATGTGTGTCAGAGT     | 4099 |
| QY | 1406 | tgtatgtatgaacactctctcttcccttatccaatctgcgtcaggtgtgccagaagcatccta      | 1465 |



```
|||||
Db 4100 TTGATGATGACACACTCTTCCTTATCCAAATTCGCTCAGTTGCCAAGAAAGATCCTA 4159
|||
Qy 1466 aaacttggfatacatcaatctgctgctgaagaggagactggactatgcttccttagtc 1525
|||
Db 4160 AACTTTGGGTACATTACATTTGCTCTGAAGAGGAGACTGGACTATGCTCCCTTAGTCC 4219
|||
Qy 1526 tcgccccgatgacagaagaattataaaagtcataattgaaagaatggccctgaaggattg 1585
|||||
Db 4220 TCGCCCCGATGACAGAAATTATAAAAGTCAATTTTGAACAAAGGCCCTCAGCGGATGG 4279
|||
Qy 1586 gtagaagatcaaaaaagctccgactttagcatatgcatcacacagatgaaaccttlaagactg 1645
|||||
Db 4280 GTAGAGAGTACAAAAGTCCGATTTATGGATACACAGATGAAGAACTTTAAGCTGCTG 4339
|||
Qy 1446 aagctatcagcatgaaatcaggaactctggagacttactttaggggaagttagagaca 1705
|||||
Db 4340 AAGCTATTCAGCATGAATTCAGGAATCTTGGGACCTTACCTTATGGGGAAGTTGAGACA 4399
|||
Qy 1706 cactgtgatattataattaagaatcaagaagcagacatatacatctacccctcagaa 1765
|||||
Db 4400 CACTGTGATTAATTAATTAAGATACAGCAAGCAACCATATACATCTACCCCTACGGAA 4459
|||
Qy 1766 tcactgactgctcctcttctatcaaggagattaccaaaagtgataaacaatttgaag 1825
|||||
Db 4460 TCACGTGATGTCCTCTTTGATTCAGAGAGATTACCAAAAAGTGTAACATTTGAAG 4519
|||
Qy 1826 atttccaatctgctcagagaaatattcaataataatggacagtgaactgtagaagt 1885
|||||
Db 4520 ATTTTCCAAATTCCTCCAGGAGAAATATTCAAATTAATGACAGTACGTAAAGATG 4579
|||
Qy 1886 ggcacaataatcaagatccctcgctgctgacccgctacttactgaatttgcataatg 1945
|||||
Db 4580 GGGCAACTTAATTCAGATTCCTTGCTGCTGACCCGCTATTAATCTGATTAATATGG 4639
|||
Qy 1946 agagagatctagcttcaagactcaatlggacctcctccatctgctataaagaactgtag 2005
|||||
Db 4640 AGAGAGATCTAGCTTCAGGACTGCTGGCCCTCTCCTCATCTGCTACAAAGAACTGTAG 4699
|||
Qy 2006 atcaagaaggaacaacagataatgtcagaacaaggaatgtcaatcctgttcttctgattg 2065
|||||
Db 4700 ATCAAAAGAGAAACACATATATGTCAACAAAGAGAAATGTCATCTCTTCTATTTTG 4759
|||
Qy 2066 atgagaacggaagctgtgtactctcacagagaatatataacgcttctcccaatccagt 2125
|||||
Db 4760 ATGGAACCCGAAGCTGTACTCTCACAGAAATATACAAAGCTTTCTCCCAATCCAGCTG 4819
|||
Qy 2126 gaatgacagcttgaagatccagagatcccaagcctcccaacatcatgacagcatatgct 2185
|||||
Db 4820 GAGTGCAGCTTGAAGATCCAGAGTCCAAAGCTCCAAACATCATGCACAGCATCATGGCT 4879
|||
Qy 2186 atgttttcttagtcttgcagttgtcagttgttctgtagaagtgtagaactgatactc 2245
|||||
Db 4880 ATGTTTTGATAGTTTGCACATTTGTCATTTGTTGTCATGAGTGTCATCTGATGATTC 4939
|||
Qy 2246 taagcatctggaagcagacagactcctctctgtctctctctcctggaatataccctaac 2305
|||||
Db 4940 TAAACATTTGAGACACAGACTGACTCTTCTTCTGTCTTCTTCTGATATACCTTCAAC 4999
|||
Qy 2306 acaaatgtctatgaaagacacatcccatatcccatctcagagaaactgtctca 2365
|||||
Db 5000 ACAAAATGCTCTATGAAAGACACATCACCCATTTCCATTCCTCAGAGAAAGTGTTCA 5059
|||
Qy 2366 tctcgatggaanaaaccaagctctatgatactctgggtgtgcacaactcaagacttcgaga 2425
|||||
Db 5060 TGTGATGAGAAAACCCAGGCTCTATGATTCGTGGGTGTCACAACTCAGACTTTCGGAACA 5119
|||
Qy 2426 gaagcatgacgccttactgaaggttctagttgtgacaagaacacgtggtattatagg 2485
|||||
Db 5120 GAGGCAATGACCGCTTACTAAAGGTTTCTAGTTGTGACAAAGAACACTGGGATTAATTAAG 5179
|||
Qy 2486 aggaacgttataagaatatctcagcatactgtcgaataaacaatgcatgaaacaa 2545
|||||

|||||
Db 5180 AGGACAGTTATGAAGATATTTACAGCATACTGCTGAGTAAAAACATNGCATGACCAA 5239
|||
Qy 2546 gaagcttctcccaagaattcaagaacccctagcactgagcgaagaatataatgccacc 2605
|||||
Db 5240 GAACTTCTCCAGAAATTCAGACACCCCTACACTAGGCAAAACAAATTTAATGCCACC 5299
|||
Qy 2606 caccagcttgaaaacgcacatcaacgggaataaactcgttactacactctcagccaatag 2665
|||||
Db 5300 CACCAGTCTTGAAGCGCATACAGGGAATTAATCTCGTACTACTCTTCAGTCAAGTCAAG 5359
|||
Qy 2666 aggaaatctgactatgatalcaccatcatcagtttgaatgaagaagaagatttgcatt 2725
|||||
Db 5360 AGGAAATGATATGATGATATCCATATCAGTTGAATGAATGAAGAAAGATTTGACATTT 5419
|||
Qy 2726 atgataaggaatgaatcagaagcccgcaagcttccaagaagaacacagacataattta 2785
|||||
Db 5420 ATGATGAGAGTAAATCAGAGCCCCCGACACTTTCAAAAGAAAACAGACACTATTTTA 5479
|||
Qy 2786 ttgctgcaatggagaagctctgggattatggatgagtagtccccaatgttcaagaa 2845
|||||
Db 5480 TTGCTGCAAGTGAAGAGCTCTGGGATTAATGGATGAGTACTCCCAATCTTTAAGAA 5539
|||
Qy 2846 acagagctcagagtgagcagtgctccctcagttcaagaaagttgtttccaagaaattactg 2905
|||||
Db 5540 ACAGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTCCAGGAATTTACTG 5599
|||
Qy 2906 atggctccttactcagcccttataccgttgagagaactaatgaaactttggactcctg 2965
|||||
Db 5600 ATGGCTCTTACTACACCCCTTAATACCGTGAAGCACTAAATGAACATTTGGGACTCTCG 5659
|||
Qy 2966 ggcacatataagagcagaagttagaataatatacatcagtgtaactttaaagaatagcct 3025
|||||
Db 5660 GGCATTAATAGAGCAGAACTTAAGATTAATATCATGATGTAATCTTGAATATAGGCT 5719
|||
Qy 3026 ctgctccctatctccttacttactagccttattcttlaatgagaaatcagagcaagag 3085
|||||
Db 5720 CTGCTCCCTATTCCTTCTATTTCTAGCTTAATTTCTTAAGGAAGATTCAGAGCAAGAG 5779
|||
Qy 3086 cagaacctagaanaaacttgtcagaagccttaatgaacccaaaacttactttagaagtgc 3145
|||||
Db 5780 CAGAACCTAGAAAAAATTTGTCAAGCCTAATGAAACCAAAATCTTATTTGGAAAGTGC 5839
|||
Qy 3146 aacatcatgagcaccacaaagaatagtttgactctcagaagccttggtcttattctctg 3205
|||||
Db 5840 AACATCTAATGCGACCCCATTAAGATGATGATGTGCTCAAAAGCTGGGCTTATTTCTTG 5899
|||
Qy 3206 atgtgacctggaanaaagatgtgaccccaagcctgcatctggaacccctctgctgcaca 3265
|||||
Db 5900 ATGTTGACCTGGAAGAAAGATGTGACATCAAGGCTGATTTGAGCCCTCTGTGCGCACA 5959
|||
Qy 3266 ctaaacactggaacccctgtctcaatggagaacaagtgtacagtaacagaaattgtctgttc 3325
|||||
Db 5960 CTAAACACACTGAACCCGTGCTATGGAGACAAAGTACAGAAATTTGCTGTGTTTT 6019
|||
Qy 3326 tcaaccatcttgaatgagaccaaagctggtgactcactcactgtaaaatagaagaaagacgca 3385
|||||
Db 6020 TCACATCTTTGATGAGACCAAAAAGCTGTTACTTCACTGTAATAATATGGAAGAAACTGCA 6079
|||
Qy 3386 gggctccctgcaatalccagaatggaagatcccactttaaagaatlatatcgtctccatg 3445
|||||
Db 6080 GGGCTCCCTGCATATATCCAGATGGAAGATCCCACTTTTAAAGAAATTAATTCGTTCCATG 6139
|||
Qy 3446 caatcaatggctacataatgataactactcgtttagtaatggtcagagatcaaaaga 3505
|||||
Db 6140 CAATCAATGCTCTACATATATGATACACTTACCTGGCTTAGTAATGGCTCAGAGATCAAAAGA 6199
|||
Qy 3506 ttcgatgtatctgctcaagatggcagcagaatgaanaacatcacttcatcatctttagtg 3565
|||||
Db 6200 TTGAGTGTATTCCTCAAGATGGGACGAATGAAGAAACATTCATTTATTCATTTCAAGTG 6259
|||
Qy 3566 gacatgttctcactgtlacgaaaaaagaagagtlataaaatggaactgttacaatctcctac 3625
|||||
Db 6260 GACATGTCTCACGTGTAAGAAAAAGAGAGATTAATAATGGCACTGTACAAATCTGTATC 6319
|||||
```

3626 caaggcttcttgagacagtggaatagttaccatcccaagcttggaattctggcgggtggaat 3685  
|||||  
6320 CAGGTGTTTTGAGACAGTGGAAATGTTACCATCCAAAGCTGGAAFTTGGCGGGGGAAT 6379  
|||||  
3686 gacctatctgacagcatctacatgctggatgagcacactttctggtgtacagcaata 3745  
|||||  
6380 GCCTTATTTGGGAGACATCTACATGCTGGGATGAGCACACTTTTCTGGTGTACACCAATA 6439  
|||||  
3746 agtctcaagaccctctgggaatgctcttgacacatctagaagattctcaagattcaagctt 3805  
|||||  
6440 AGTCTCAGACTCCCTGGGAATGGCTTTGACACATTAGAGATTTTTCAGATTACACTT 6499  
|||||  
3806 caagacaataatgagacagtgggcccccaaaagctggcagacattcatatctccgattcaatca 3865  
|||||  
6500 CAGSACATATGAGACAGTGGGCCCCCAAGCTGGCCAGACTTCATTATTCGGATCAATCA 6559  
|||||  
3866 atgctctggagacacaaaggaagcccttctcttgatcaaaagtgtgactctgttgacacaaatga 3925  
|||||  
6560 ATGCTTGAGAGCACCAGAGGCCCTTTCTTGATCAGAGTGATGCTGTGGCACCACCAATGA 6619  
|||||  
3926 ttattcaagcatcaagaaccccaagggtgccgttcagaagttctccagaccctacatctctc 3985  
|||||  
6620 TTATTCACGGCATTAAGACCCAGGGTGCCCTGACAGACTTCTCCAGCCTTCAATCTCTC 6679  
|||||  
3986 agttatcatcatgtatagttcttgatggaaagtggcagactatcgaaggaattcca 4045  
|||||  
6680 AGTTATCATCATGTATAGTCTTGATGGGAAGAAAGTGACAGACTATATCGAGGAATTTCCA 6739  
|||||  
4046 ctggaaccttaatggtctctcttggaatgtggtatcaatctgggaataaacaataattt 4105  
|||||  
6740 CTGGAACCTTAATAGTCTCTTTTGGCAATGGATTCATCTGGGATTAACCAATATTT 6799  
|||||  
4106 ttacccttccaatatttgctgcatatcatccgtttgaccaccaatcatatagatctgca 4165  
|||||  
6800 TTAACTCCATATTATGCTGTGATACATCCGTTTGACCCAACTATATATACATTTCCGA 6859  
|||||  
4166 gcaactctgcataagagttgaatggcctgtgaatttaaatagttgcagcaatggcattgggaa 4225  
|||||  
6860 GCACCTTTCGCATGAGAGTGTATGGGCTGTGATTTAAATAGTTGACGATGGCATTTGGGAA 6919  
|||||  
4226 tggagaatgaagaacatatcagaatgacagattactgcttcaatctactacttaccataatgt 4285  
|||||  
6920 TGGAGATTAAGCAATATACATGACAGATTACTGCTTCACTTCCATCTTACCAATATGT 6979  
|||||  
4286 ttgcacctgtgtctccttcaaaaagctgacttcaacttcaagggaggaagtatgtctgga 4345  
|||||  
6980 TTGGCAGCTGCTCTCTTCAAAAGCTGCACCTTCCAAAGGAGGAGTATGCTTGGGA 7039  
|||||  
4346 gaacctgagttgaataatccaaaagagtggtgcaagtgagacttccagaagaacatgaaag 4405  
|||||  
7040 GACCTCAGGTGAATTAATCCAAAGAGTGGCTGCAAGTGGACTTCCAGAAAGCAATGAAG 7099  
|||||  
4406 tccagaagaatgaactactcagaaggaataaaattctgcttcaacagcaatgatatgtgaagagt 4465  
|||||  
7100 TCACAGGAGTAACATCTCAGGAGTAATAATCTCTGCTTACACAGCAATGATGGAAGAGT 7159  
|||||  
4466 tcccatcctcaagcagtcgaagatggccaatcagtgagctctcttctcaagaatggcgaag 4525  
|||||  
7160 TCCTCATCTCCAGAGATCAAGATGGCATTCAGTGACCTCTTTTTCAGAAATGGCAAG 7219  
|||||  
4526 taaaagttttcagggaataatcaagactcctcacaacctgtgtgaaactctcagaacccac 4585  
|||||  
7220 TAAAGGTTTTTTCAGGAAATCAAGACTCTTACACACTGTGTGAATCTCTAGACCCAC 7279  
|||||  
4586 cgttactgactcgtctaccctcgaatcaacccccagaagtgtgtgacacagatgtccctga 4645  
|||||  
7280 CGTTTACTGACTCGCTACTCTTGAATTCACCCCAAGAGTTGGGTGACCAAGATTTGCCCTGA 7339  
|||||  
4646 ggaatgaggtctcgggctcgaaggacaggaaccttactagctcgaagagttc 4699  
|||||  
7340 GGATGAGGTTTCTGGGCTGCGAGCACAGACCTCTACTGAGGTGGCCACTGC 7393  
|||||

RESULT 2  
AR110040 4670 bp DNA PAT 14-FEB-2001  
LOCUS AR110040  
DEFINITION Sequence 41 from patent US 6114148.  
ACCESSION AR110040  
VERSION AR110040.1 GI:12826316  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE  
1 (bases 1 to 4670)  
AUTHORS Seed, B. and Haas, J.  
TITLE High level expression of proteins  
JOURNAL Patent: US 6114148-A 41 05-SEP-2000;  
FEATURES  
LOCATION/Qualifiers  
SOURCE 1..4670  
BASE COUNT 1325 a 1072 c 1036 g 1237 t  
ORIGIN  
Query Match 91.5% Score 4357.6; DB 9; Length 4670;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 4363; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
328 gccaccagaagaatactacctgggtgacgtggaactgcatgtggaactatagcaatgcat 387  
|||||  
84 GCCACCAGAGACTACTACTGCTGCTGACAGTGTGCTGACAGTGTATGGAATATGCAAGTAT 143  
|||||  
388 ctggtgacgtgctgtggaacgaatcttccctagatgacaaatcttccattc 447  
|||||  
144 CTCGCTGAGCTGCCCTGTGACGACGAAGATTTCCTCTAGAGCCAAATCTTTTCATTC 203  
|||||  
448 aacacactaagtcgtgtacaaaaagactctgtttttagaatcaaggttcaacttltaac 507  
|||||  
204 AACACCTCAGTCTGTATCAAAAAAGACTCTGTTTATGAAATCAAGGATACACCTTTTCAAC 263  
|||||  
508 atgcgttaagcaaaagccaccctgtagtggctgtctgcttaccatcccaagcgaagtc 567  
|||||  
264 ATCGCTAAGCCAAAGGCCACCTGGATGGGTGCTGCTAGTCTTACCATCCAGGCTGAGGTT 323  
|||||  
568 tatgataagtgatcatctacattaaagacatggcttccatctcgttcaatcttcaatgct 627  
|||||  
324 TATGATACAGTGTATACACTTAAGAACATGGCTTCCCTGTCAGACTTTCATGCT 383  
|||||  
628 gttgtgatactctactcgtgaaagctctctgaggaactgaataatgataatgacagcaacaa 687  
|||||  
384 GTTGTGTATCTCTACTGGAAGCTTCTGAGGAGCTGAAATATGATATGATACAGACGTCAA 443  
|||||  
688 agggagaagaagaatgataaagctctccctgtgtgaaagccatataatgctgcaagtc 747  
|||||  
444 AGGGAGAAAGAAATGATTAAGTCTTCCCTGTGTGAAGCCATATATATGCTGGCAGTTC 503  
|||||  
748 ctgaaagaagaatgtgccaatgagctctgagccactgtgcttactactcatatcttctc 807  
|||||  
504 CTGAAGAGAAATGTGCTCATGAGCTGACCCACTGTGCTTACTACTCATATATCTTCT 563  
|||||  
808 catgtgagccgtgtgaaagacttgaatcagcctcctcaatcttgaagccctactagatgaga 867  
|||||  
564 CATGTGACCTGGTAAAGACTTGAAATTCAGGCTCATGTGAGGCCCTACTAGATATGTA 623  
|||||  
868 gaaggagactctggccaaagaaagacacagaccccttgacaaattcatctactcttctc 927  
|||||  
624 GAAGGAGTCTGGCCAAAGAAAGACACAGACCTTGCACAAATTTATATACACTTTTGTCT 683  
|||||  
928 gtatttgaatgaaggaaagtgtggaactcagaagaacaaagaactcctgtatgcaagatgag 987  
|||||  
684 GTATTTGATGAAGGAAAGTTGGCACCTCAGAAACAAAGAACTCTGTATGACAGATAGG 743  
|||||  
988 gatgtctgatactgctcgggctcgtgctcaaaatgacacacagtcaatggttatgtaaagag 1047  
|||||  
744 GATGCTGATCTGCTGGGCTTGGCTTAAATGCAACAGTCAATGATGTTATGTAACAGG 803  
|||||

|    |      |  |      |
|----|------|--|------|
| QY | 1048 | tctctccaggtcttgattggaatgcacaggaatcaatctatctggcatgtgattggaatg     | 1107 |
| Db | 804  | TCCTCGCAGAGGTCTGATTTGGATGTCACAGAAATCAGCTTATTTGGCATGTGATTTGGATG   | 863  |
| QY | 1108 | ggcaccaactccttgaagtgactcaataltcctcgaagggtcacacatctctgttgaaggac   | 1167 |
| Db | 864  | GGCACACACTCTCGAAGTGACACTCAATATTTCTCGAAGGTACACATTTCTTGAGGAGAC     | 923  |
| QY | 1168 | catgcgcaagggtctcttggaaatctgcgaataactcttctactgtctcaaacctcttg      | 1227 |
| Db | 924  | CATGCGCAGGGGCTCTGGAAATCTCGCAATTAACCTTCTACTGCTCAAAACCTCTTG        | 983  |
| QY | 1228 | atggaccttgagcaagttctactctgtttgtacatactctctcccaacaatgagatgagatg   | 1287 |
| Db | 984  | ATGGACCTTGGACAGATTCTACTCTGTTTGTATCTCTTCCCAACCATGATGGCATG         | 1043 |
| QY | 1288 | gaagcttatgtcacaagtagaagcgtgtccagaggaaccccaactacgaatgaaaaataat    | 1347 |
| Db | 1044 | GAAGCTTATGTCAAAAGTAGACAGCTGTCCAGAGAGAACCCCAACTACGAATGAAAAATAT    | 1103 |
| QY | 1348 | gaagaagcgggaagactatgatgatctctactgtacttgaatgtgattgtgcaagttc       | 1407 |
| Db | 1104 | GAAAGAGCGGAAGACTATGATGATGATCTTACTGATTCTGAAATGATGTGGTCAGGCTTT     | 1163 |
| QY | 1408 | gatgatgaacaactctctctctcttatacacaattgcgcagttgcgaagaaagcatcctaaa   | 1467 |
| Db | 1164 | GATGATGACACACTCTCTCTCTTATCCAAATTTGCTGCTAGTTCGCAAGAAACCATCTPAA    | 1223 |
| QY | 1468 | acttgggtacatllacatctgtcgtcgaagaggaagcgtggaactatgtctccctatgtctc   | 1527 |
| Db | 1224 | ACTTGGGTACATTTACATTTGGCTGGAAGAGAGACTGGGACTGTGCTCCCTTAGTCTCTC     | 1283 |
| QY | 1528 | ggccccgatgacagaagttataaaagtcaatattgaaacaatggccctcagcggattggt     | 1587 |
| Db | 1284 | GCCCCCATGACAGAGTTATTAATAAGTCAAATTTGAAACAATGGCCCTCAGCGGATTTGCT    | 1343 |
| QY | 1588 | aggaagtaacaaaagtcgcgaattatggacacacagatgaacactttaagaccctgaa       | 1647 |
| Db | 1344 | AGGAAGTACAAAAGATCCGATTTATGGCATACACAGTGAACCTTTAAGACTCGTGA         | 1403 |
| QY | 1648 | gcatllacgaatgaatcaagaatcttggaccttacttlatatggggaagtgtggaacaca     | 1707 |
| Db | 1404 | GCTATTGAGCATGATGATGAGGATCTTGGACCTTACTTATGSGGAATTTGAGACACA        | 1463 |
| QY | 1708 | ctgtgtgattatatttaagaatcaagcaagcagacacataacatctacccctcaaggaatc    | 1767 |
| Db | 1464 | CTGTGTGATTTATTTAAGAAATCAAGCAAGACACATATTAACATCTACCCCTCAGGGAATC    | 1523 |
| QY | 1768 | actgatctcgtctcttctgtatllcaagagagatllaccaaaaggtgaaacaatllgaagat   | 1827 |
| Db | 1524 | ACTGATCTCGCTCTCTTTGTATTTCAAGAGATTTACCAAAAGGTGTAACATTTTGAAGAT     | 1583 |
| QY | 1828 | tttccaatctcgcagagagaataatllcaaatataatgaagcagtgagctgtagaagatg     | 1887 |
| Db | 1584 | TTTCCAAATTTGCCAGAGAGAAATATTCAAATATTAATGGACAGTGTGAAGATGGG         | 1643 |
| QY | 1888 | ccaactaatlcaagatcctctgtgtgcctgaacccgctataactctagtttcgttaataatg   | 1947 |
| Db | 1644 | CCAACCTAAATCTAGATCTCGGTGCTTACCCGCTATTACTCTAATTTCTGTTAAATATGAG    | 1703 |
| QY | 1948 | agagatctagctcaggaactatgtgcctctctctcactcgtcgaacaagaatcgttagat     | 2007 |
| Db | 1704 | AGAGATCTAGCTTACAGACATCATTTGGCCCTCTCTCATCTGCTCAAAAGAAATCTGTAGAT   | 1763 |
| QY | 2008 | caaaaggaanaacagataatgtcagacaagaggaatgtlcalcctgtttctgtlatgtat     | 2067 |
| Db | 1764 | CAAAAGGAAACCCAGATTAATGTCAAGACAAGAGAAATGTCAATCTGTTCTGTGATTTGAT    | 1823 |
| QY | 2068 | gagaacccgaagctgtgtaccccaagagaaatatacaagctttctcccaatccagctgga     | 2127 |
| Db | 1824 | GAGAACCGAAAGCTGTGCTCTACAGAGAAATATACAAAGCTTTCTCCCAATCCAGCTGGA     | 1883 |
| QY | 2128 | gtgcagcttgaggaatccagagttccaaagcctccaacatcatatgcacagcatcaatgtctat | 2187 |
| Db | 1884 | GTGCAGCTTGAGGATCCAGAGATTTCCAAAGCTTCCAACTCATGTGACACATCAATGCTTAT   | 1943 |
| QY | 2188 | gtttttagatagtttgcaggtgtcagttgtgtttgtatgaaggtgtgcatatctgtatctia   | 2247 |
| Db | 1944 | GTTTTGTATAGTTTGCAGTTGTCTAGTTTGTGTTGTGATGAGGTGGCATCTGTATATTCTA    | 2003 |
| QY | 2248 | agcaatggagacagactgacttctctctgtctctctctctctctctctctctctctctct     | 2307 |
| Db | 2004 | AGCATTTGGACACAGACTGACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT    | 2063 |
| QY | 2308 | aaaatgtctatgaagacaacactcaacctatctccactctcccaagagaacatgtcttcatg   | 2367 |
| Db | 2064 | AAAATGCTTATGAGACACACACTCACCTTATCCCATTTCTCAGAGAAACTGTCTTATG       | 2123 |
| QY | 2368 | tcgaatgaaaaacccaggtctatagatcttgyggtgcacaactcaagacttctggaacaga    | 2427 |
| Db | 2124 | TCGATGGAAAACCCAGGTCTATGATTTCTGGGGTCCCAACACTCAGACTTTCGGAACAGA     | 2183 |
| QY | 2428 | ggcatgacgcgtctactcgaaggttctcaagtgtgtgaagaagacactgtgtattatcagag   | 2487 |
| Db | 2184 | GGCATGACCCGCTTACHTGAAGGTTTCTAGTTGTGACAAGAAACACTGGTGAATTTACGAG    | 2243 |
| QY | 2488 | gacagttatgaagatatttcaagcaacttctgtgtgtgaataaacaatgacattgaaccaga   | 2547 |
| Db | 2244 | GACAGTTATGAGATATTTTACGACTTCTGCTGATGTAACAAATGCGATTTGAACACAGA      | 2303 |
| QY | 2548 | agcttctccagaatllcaagaacacccctagacactaggggaanaagcaattlaatgcccacca | 2607 |
| Db | 2304 | AGCTTCTCCAGAAATTCAGACACACCTTAGCATGAGCAAAAGCAATTAATGGCACCCCA      | 2363 |
| QY | 2608 | ccagctcttgaagcgcatcaacgggaaataactctgtactactctctcagtaagaacagag    | 2667 |
| Db | 2364 | CCAGCTTTGAAACCCCATCAACGGGAATTAACCTGCTCTCTCTCTCTCTCTCTCTCTCTCT    | 2423 |
| QY | 2668 | gaatltgacctatgataacataatcagttgaagatgaagaaggaagatlltgacattat      | 2727 |
| Db | 2424 | GAAATTTACTATGATGATATACCATATATCTGTTGAATGAAAGGAAGATTTTGGACATTTAT   | 2483 |
| QY | 2728 | gatgaagatgaanaalccagagcccccgagctllcaaaaagaanaacagacacatattat     | 2787 |
| Db | 2484 | GATGAGGATGAATAATTCAGACCCCCCGCAGCTTCAAAAAGAAACACAGACTATTTTAT      | 2543 |
| QY | 2788 | gtctgagatgagaggtcctctgtgtatattgagatgtgtgtgtgtgtgtgtgtgtgtgtgt    | 2847 |
| Db | 2544 | GCTGCAGTGAAGAGGCTCTGGGATTAATGAGATGAGTGTGCTCCCACTGTCTTAAGAAC      | 2603 |
| QY | 2848 | agggctcagagtgagctgtccctcagttcaagaagaatgttttccaggaatttaactgat     | 2907 |
| Db | 2604 | AGGGCTCAGAGTGGCAGTGTCCCTCACTTCAAGAAATTTGTTTCCAGGAATTTACTGAT      | 2663 |
| QY | 2908 | ggtctcttactcagccccctataccgtlgaagaactaaatgaacatltggaactcctlgyg    | 2967 |
| Db | 2664 | GGCTCTTACTACTGACCCCTTATACCGTGAGAACTAAATGAACATTTGGGACTCTG6GG      | 2723 |
| QY | 2968 | ccatataatgaagcagaagtttgaagataataatcatgttlaactcttcaagaatcagggcctc | 3027 |
| Db | 2724 | CCATATATTAAGAGCAGAGTGAAGATTAATCATGTGTAATCTTCAAAATTCAGGCTCT       | 2783 |
| QY | 3028 | cgctccatctctctctatctttagactatcttctatgaagaagaatcagaagcaagagca     | 3087 |
| Db | 2784 | CGTCCCTATTTCTTCTATTTAGCTTATTTCTTATTTATGAGAAATCTCAGAGCAAGAGCA     | 2843 |
| QY | 3088 | gaacctagaanaaaactltgtcaagcctaaatgaacccaaaacttactctltggaagtgc     | 3147 |
| Db | 2844 | GAACCTGAAAAAATTTGTGCAAGCCTTAATGAACCAAAACCTTACTTTGGAAGTGC         | 2903 |
| QY | 3148 | catcataatgaccccaataaagatgtgtgactgtcaagcctgggcttatctctctgat       | 3207 |
| Db | 2904 | CATCATATGGCACCCCACTAAAGATGAATTTGACTGCAAAACCTGGGGCTTATTTCTCTGAT   | 2963 |
| QY | 3208 | gttgacctggaanaagatgtgcactcagcctgtatgtgaacccctctgtctgtccaaact     | 3267 |

|||||  
Db 2964 GTTACCGGAAAAAGATGTGCACACGAGCCGATGTGACCCCTCTGTGTCACACACT 3023  
Oy 3268 aacacactgaacccctgctcatcgggagacagtgacagtaacagaaatttgctgttttc 3327  
Db 3024 AACACACTGAACCCCTGCTCATGCGGAGACAGTACAGAGTAATTTGCTGTGTTTTTC 3083  
Oy 3328 accactcttgaatgaacacaaagctgtaactcactgaaataatgaaagaaactgcag 3387  
Db 3084 ACCATCTTGTATGAGACCAAAAGCTGTACTTCACTGAAATATGAAAGAACTGCAGG 3143  
Oy 3388 gctccctgcaatatccagatggaagatccacatttaagagaattatcgcttccatgca 3447  
Db 3144 GCTCCGCAATATCCAGATGGAGATCCCACTTTAAAGAAATATTCGCTTCATGCA 3203  
Oy 3448 atcaatggtcctaataatgataactactgctgctgaatgaatggtcgaagataaaggatt 3507  
Db 3204 ATCAATGGCTACATATATGATACACTACCTGGCTTATGATGCTGAGATCAAAAGATT 3263  
Oy 3508 cgaatgatactgctcagacatggtgcagcaatgaanaacatccatctatccatcagtgca 3567  
Db 3264 CGATGATATCTGCTCAGCATGGGAGCAATGAAACATCCATTTATTCATTTCAGTGA 3323  
Oy 3568 catggtcactgtaacgaanaaagaagatgaataatggaactgtaacatctcaca 3627  
Db 3324 CATGTGTTCACTGTACGAAAAAAGAGATATAAATGCGACTACATCTTATCA 3383  
Oy 3628 ggtgtttttgagacagtggaatgttaccatccaaagctggaattgacgggtggaatgc 3687  
Db 3384 GGTGTTTTTGAGACAGTGGAAATGTTACCATCCAAAGCTGGAAATTTGGCGGTGGAAATGC 3443  
Oy 3688 ctatctgagcagcatctcaatgctggaatgagacacatttctggtgtacagcaataag 3747  
Db 3444 CTATATGGCGAGCATCTCATGTGCTGGATGAGCACACTTTTCTGCTGACAGCAATAAG 3503  
Oy 3748 tgttaagactccccggggaatggtctcttgagacatataagatttccagattacagcttca 3807  
Db 3504 TGTGAGACTCCCTGGGAAATGGCTTCTGACACATTAAGATTTCAGATTACACCTTCA 3563  
Oy 3808 ggaacataatgacagatggtgccccaaagcttgccagacttcatctatccgagcaatcaat 3867  
Db 3564 GGAACAATATGACAGATGGGCCCCAAAGCTGGCCAGACTTATTCGGAGCAATCAAT 3623  
Oy 3868 gctctgagcaccaaaggaccttctcttgatcaaatgtaabctgttgacacaaatgatt 3927  
Db 3624 GCGTGGAGACACAAAGAGCCCTTTCTTGATCAAGTGTGATCTGTGGACCAATGATT 3683  
Oy 3928 attcaaggatcaagaacccaaggtgtccgtcaagaagttctcagctctacatctcag 3987  
Db 3684 ATTCACGGCATCAAGACCCAGGTCGCCGTGAGAAATCTCCAGCCTTACATCTCTCAG 3743  
Oy 3988 tttaatacatgataagcttgatgaggaagatgagagacttccgaaggaaatccact 4047  
Db 3744 TTTATCATCATGATAGTCTTGATGGAAGAAAGTGGACACTTATTCGAGAAATTCACAT 3803  
Oy 4048 ggaacctaataatgctctctcttgcaaatgtaattcatctggaataaacaataattctt 4107  
Db 3804 GGAACCTTAATGTCTTCTTGTGGCAATGTGATTCATCTGGGATTAACACAAATATTCTT 3863  
Oy 4108 aacacctcaatatctgctcgataatccgcttgtaacccaactcatatagacattgcagc 4167  
Db 3864 AACCTCAATTAATGTCTGATACATCCGTTTGACCCAACTCAATTAAGCATTTGCGAGC 3923  
Oy 4168 accttcgacatgagatgagctggtgatttaaatagttcgaacatgcaattggaattg 4227  
Db 3924 ACTTGTGCAATGAGATGATGGCTGTGATTTAAATAGTTGAGCATGTCATTTGGGAATG 3983  
Oy 4228 gagaatgaagaatatacagaatgacagataatcgtcttaccatcttaccataatgttt 4287  
Db 3984 GAGAGTAAGCAATATCATGATGACAGATTAATCTGCTCATCTTACCAATATGTTT 4043  
Oy 4288 gcaacactggtctccttcaaaagctcgacttcaacctccaaggaggaagatgacctggaga 4347  
|||||

Db 4044 GCCACCTGGTCTCTTCAAAAAGCTGCACCTTCACAGGAGGAGTATGCTTGAGA 4103  
Oy 4348 cctcagatgaatataatccaaagaagatgctgcaagtgcagcttccagaagaatgaagtc 4407  
Db 4104 CCTCAGGTGAATATATCCAAAGAGTGGCTGCAAGTGCACCTTCAGAAACATGAATGC 4163  
Oy 4408 acagaatgaactactcaaggagtaaatctctgcttaccagaatgtaatgaaagattc 4467  
Db 4164 ACAGAGTAATCACTACAGGAGTAAATCTGCTTACCAACATGATGTAAGAGACTTC 4223  
Oy 4468 ctcatctcagagatcagaatagggcatagtggtctcttcttccagaatgaagta 4527  
Db 4224 CTTCATCTCAGAGTCAAGATGAGCCATCATGATGATCTCTTTTTCAGAAATGCAAGTA 4283  
Oy 4528 aaggttttccaggaatcaaacactctcaaacactggtggtgaactctcagaccacg 4587  
Db 4284 AAGGTTTTTCAAGGAATATCAAGACTCTTCACACTGTGGTGAACCTCTAGACCCACG 4343  
Oy 4588 ttactgactcgtactactcgaattccaccacagattggtgacacagatgacctgag 4647  
Db 4344 TTACTGACTCGCTACTCTTCGAATTCACCCCGAGTGGTGACACAGATTGCCCTGAGG 4403  
Oy 4648 atggaagttctggtgctgcgaggaagaagactctcactgactcagagcgagttc 4699  
Db 4404 ATGAGAGTTCTGGGCTGGAGCACAGACCTTACTGAGGGTGGCCACTGC 4455  
  
RESULT 3  
108644  
LOCUS 108644 4278 bp PAT 02-DEC-1994  
DEFINITION Sequence 4 from Patent WO 8800831.  
ACCESSION 108644  
VERSION 108644.1 GI:588650  
KEYWORDS .  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4278)  
AUTHORS Pasek, M.P.  
TITLE DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE  
POLYPEPTIDES IN HIGH YIELDS  
JOURNAL Patent: WO 8800831-A 4 11-FEB-1988;  
FEATURES  
source 1..4278  
BASE COUNT 1244 a 942 c 946 g 1146 t  
ORIGIN  
  
Query Match 87.1%; Score 4148.8; DB 10; Length 4278;  
Best Local Similarity 97.9%; Pred. No. 0;  
Matches 4270; Conservative 0; Mismatches 2; Indels 90; Gaps 3;  
  
Oy 328 gccacagaagaatcactacccgtggtgacgtggaactgtaatgagacatataatgaatgatt 387  
Db 1 GCCACAGAAATATCTACTGCTGGTGCAGTGAACGTGCATGGCTATATGCAAAAGTAT 60  
Oy 388 ctgggtgagctggtcctggtgagcaagaatttctctgtagagtgagcaaaatctttccatt 447  
Db 61 CTCGGTGAAGCTGCTGTGACGCAAGATTTCTCTAGTGAAGTGCAGAAATCTTTTCATTC 120  
Oy 448 aacacctcagtcgtgtacaaaagaactctgtttgataatcaagttcacacttltcaac 507  
Db 121 AACACCTCAGTCGTGTACAAAAGACTCTGTGTGATATCAAGATCAACCTTTTCAC 180  
Oy 508 atcgttaagcaaggtcgaacctggtatggtctgtctaggttctaccatcgaagctgagtt 567  
Db 181 ATCGCTAAGCCAAAGGCCCTGATGGGTCTGCTAGTCTTACCATTCAGGCTGAGATT 240  
Oy 568 tatgataagtggtcattacacttaagaagatgcttccatccatctgtcattcttcatgct 627  
Db 241 TATGATACAGTGTGCTATTAACCTTAAGAACATGGCTTCCACTGTGAGCTTCATGCT 300  
|||||



|||||  
Db 2377 ATTGGTGAAGTGAAGGCTGTGGGATTTATGGATGATAGTACGCCACATGTTCTAAG 2436  
2845 aacaaggctcaagatgagcagtgctccctcagttcaagaagatgtttccaggaattact 2904  
2437 AACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTTCCAGGAATTTACT 2496  
2905 gatgctccttactcaagcctcctataccgttgagaaactaataagacatttgagcctctg 2964  
2497 GATGCTCCTTACTCAGGCCCTTATACCTGGAGAACTTAATGAACATTTGGACTCTCG 2556  
2965 gggccatataaagagcagaagttgaagataataatcattgaacttcagaatcagcc 3024  
2357 GGGCCATATATAGAGCACAAGCTTGAAGATATCATCTGTAACCTTGAGAAATTCAGGCC 2616  
3025 tctcgtccctatctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 3084  
2617 TCTGTCCTCTATCTCTTCTATCTAGCCCTTATTTCTTATGAGAGACATCAGAGCAAGGA 2676  
3085 gcagaacctagaanaaacttgtcaagcctaa tgaaccanaaacttacttctgaaagt 3144  
2677 GCAGAACCTAGAAAACCTTTGTCAAGCCTAATGAACCAAAACCTTATTTGGAAGTG 2736  
3145 caacatcataatggcaccactaaagaatgatttgactgcaagaagcctgggcttattctct 3204  
2737 CAACATCATATGAGCACCACCTAAAGATGATTGACTGCAAGCCGCGCTTATTTCTCT 2796  
3205 gatgttgaccttggaanaaagtgtgacacagccttgatctggagccctctgtctgctcac 3264  
2797 GATGTTGACCTTGAAAAAGATGTGCACTAGGCCCTGATTTGGACCCCTTGTGCTCCAC 2856  
3265 actaacacactgaacccctgtcactatggagagacagttacagaaatttgctctgtt 3324  
2857 ACTAACACACTGAACCCCTGCTCATGAGAGACAAGTGACAGTACAGAAATTTGCTGT 2916  
3325 t---tcaccatcttctgtagagaccagaagctggttacttcaactggaataatggaaagaac 3381  
2917 TTCCCTGACATCTTGTGATGAGACCAAAAGCTGTACTTACGTGAATAATGTGAAAGAAC 2976  
3382 tgcagggctccctcgaatatacagaatggagaatcccaactttaaagagaattatcgttc 3441  
2977 TGCAGGGCTCCCTGCAATATTCAGATGGAAGTCCCACTTTTAAAGAGAAATTTATCCCTTC 3036  
3442 catgcaatcaatlygctacataaataatgatacactcgtgcttagtaatggctcagaatcaa 3501  
3037 CATCAATCAATGCTGATACATATGATGATACACTGCTGCTAGTATGCTCAGAGTCAA 3096  
3502 aggatctgagatctctgtcgaagcctggagcaatggaaataacatcatctctatcttctc 3561  
3097 AGGATTCAGATGATCTGCTCAGCATGGCAGCAATGAAACATCCATTCTATTCATTTTC 3156  
3562 agtggacaatgtctcactgtaagaaaaaagagagatataaaatggcaactgtacaaatc 3621  
3157 AGTGACACTGTGTTACGTACGTACGAAAAAAGAGATATAAAATGGCACTGTACAAATTC 3216  
3622 tatccaggttgttcttgagacagttgaaatgttaccatccaagctgaaatttgccggtg 3681  
3217 TATCCAGGTGTTTGTGAAGACAGTGAATGTACCATCCAAAGCTGGAATTTGGCGGGTG 3276  
3682 gaatgacctatctggagacatctacaatgtgagatgagacaccttcttctgtgtacag 3741  
3277 GAATGCCCTTATTTGGAGCATCTACATGCTGGATGAGACACTTTTGTGTTGATCAAC 3336  
3742 aataagctgagactccctggaaatggctcttgagacacatagaagatttcaagatttaca 3801  
3337 AATAAGTGTCAAGACTCCCTTGGAATGGCTTCTGACACATTTAGATTTTTCAGATTTACA 3396  
3802 gcttcagacaataatgacagtgggcccaaaagctgagccagactcattatctcagatca 3861  
3397 GCTTCAGACAATAATGACAGTGGGGCCCAAGAGCTGGCCAGACTTATTTCCGGAATCA 3456  
3662 atcaatgctgagagaccaaaggagcccttctcttgatcaaggtgagatctgttgaccca 3921  
|||||

Db 3457 ATCAATGCCGTGAGACCAAGAGACCCCTTTCTTGATCAAGTGATCTGTTGGACCA 3516  
3922 atgattattcaagcagcaatcaagaccaggttgcccgctcaagaattctcagacctacatc 3981  
3517 ATGATTATTACAGGATTCAGAACAGCCAGGTGCCCTGCAAAATTTCTCCAGCTCTACATC 3576  
3982 tctcagttatcatcatgatatagctcttgaatggaaagatggcagactatcaggaat 4041  
3577 TCTCACTTATCATCATGATGATGCTTGTGAGGAAGATGCGACAGCTTATGAGGAAT 3636  
4042 tccactggaaacttaatgtcttctcttgagcaatgagatcatctcggataaanaacat 4101  
3637 TCCACTGGAACCTTAAATGGATCTTTTGGCAATGGGATTCATCTGGGATTAACACAT 3696  
4102 attttaaccctccaaatcttgctgatacatcgtttgcaccacacatataagact 4161  
3697 ATTTTAACTTCGCAATTTTGTGATGATACCTGTTGACCCAGCACTTATTAAGCAT 3756  
4162 cgaagcactcttcgaatgagatgagtggtgctgtaattaaatagttgacagatgcaatg 4221  
3757 CCAGACACTCTTCGATGAGATGAGTGTGATGGCTGTGATTTAAATAGTTGCAGCATGGCAT 3816  
4222 ggaatggagaatgaagaacatatacagatgacagatattgcttacttacttaccat 4281  
3817 GGAATGAGAGTAAGCAATATCAGATGACAGATTACTGCTTCAATCCTACTTTACCAAT 3876  
4282 atgtttgcaactgtctctcctcctcaaaagctgcaactcctccaaaggagagatgctc 4341  
3877 ATGTTTGCACCTGTGCTCTCTTCAAAAGCTGACTTCACTCCAAAGSAGAGATATGCC 3936  
4342 tggagaacctcaagtgtaataatccaaagatgagctgcaagtgagcttccagaagaatg 4401  
3937 TGGAGACCTCAGAGTAATATCAAAAGAGTGGCTGCAAGTGGACTTCCAAAGACAAATG 3996  
4402 aaagtcaagggatcaactcactcagggagtaaatctcgtctaccagaaatgataatg 4461  
3997 AAAGTCAAGAGATCACTCAAGGAGTAATAATCTGCTTACAGCATATGATGTAAG 4056  
4462 gagtccatctcctcagcagatcaagatggcatcagtgagctctcttctcaagaatgag 4521  
4057 GAGTTCCTATCTCCAGCATCAAGATGCGCATGAGTGTGCTTTTTCAGAAATGCC 4116  
4522 aaagtaaaagtttctcagggaaatcaagaactcctcaacctgtgtgtaactctcagac 4581  
4117 AAAGTAAAGTCTTTCAGGGAATCAACACTCTTCAACCTGTGTAACCTCTAGAC 4176  
4582 ccacgttactcactcgcctcactccttgaaattcaaccccaagatgggtgaccagaatggc 4641  
4177 CCACGCTTACTGACTGTGCTTGAATTCACCCCAAGATGGGTGACCAAGATTTGCC 4236  
4642 ctgagatgagatctctgagctgagagcaagagacctcag 4683  
4237 CTGAGGATGAGGTTCTGTGGCTGCGAGCAGACGACCTCTAC 4278  
Db 4237 CTGAGGATGAGGTTCTGTGGCTGCGAGCAGACGACCTCTAC 4278

RESULT 4  
108643 4281 bp PAT 02-DEC-1994  
LOCUS 108643  
DEFINITION Sequence 3 from Patent WO 8800831.  
ACCESSION I08643  
VERSION I08643.1 GI:588649  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4281)  
AUTHORS Pasek, M.P.  
TITLE DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE  
POLYPEPTIDES IN HIGH YIELDS  
JOURNAL Patent: WO 8800831-A 3 11-FEB-1988;  
FEATURES  
source I. 4281  
Location/Qualifiers



/organism="unknown"  
BASE COUNT 1245 a 943 c 946 g 1147 t  
ORIGIN  
Query Match 87.0% Score 4147.2; DB 10; Length 4281;  
Best Local Similarity 97.9%; Pred. No. 0;  
Matches 4269; Conservative 0; Mismatches 3; Indels 90; Gaps 3;  
QY 328 gccacccagaatactactccctggtgctgagtggaactgtcagtgaggaactatagcaagagat 387  
|||||  
Db 4 GCCACCGAAGATATCTACTGCGGTGGAGTGCACTGATGGAGCTATATGCAAAAGAT 63  
QY 388 ctccgtagctgctctgtagcagcaagaattccctcctagagtgcaaaaactttccatc 447  
|||||  
Db 64 CTGGGTAGGCTGCTGTGGAGCGCAAGATTCTCTCTAGAGGCCAAATCTTTTCCATTTC 123  
QY 448 aacacccagtcgtgtacaaaagactcgtttgtagaattcaagttcacggttccattcaac 507  
|||||  
Db 124 AACACCTCATGCTGTACAAAAGACTCTGTTGTAGAATTCAAGATTCACCTTTTCAC 183  
QY 508 atgcctaagcaagggccacctgtagtggtctgtctgaagttccacatccagctgaagtc 567  
|||||  
Db 184 ATGCGTAAGCCCAAGGCCACCTGGATGGTCTGCTAGGTCCTTACCATCCAGGCTGAGGTT 243  
QY 568 tatgatacagtggtcatctacacttaagaacatggtcccatccctgfcagttcattcagtc 627  
|||||  
Db 244 TATGATTCAGTGGTCACTTAACTTAAGACATGGCTTCCATCTGTCAGTCTTTCATGCT 303  
QY 628 gttagtgatcctactgtgaaagctctctgaggagctgtaataatgatacgacacagtcacaa 687  
|||||  
Db 304 GTTGCTGTATCTACTGGAAGAGTTCTGAGGAGCTGAATATGATGATCAGACAGTCA 363  
QY 688 agggagagaagaagatgataagctctccctcgtgaggaaacatatatgtctggcaagtc 747  
|||||  
Db 364 AGGAGAGAAAGATGATTAAGTCTTCCCTGGGAGGACCATATATGCTGGCAGGTC 423  
QY 748 ctgaagaagaatggtcccaatgagcctctgacccactgtgcttacctactatattcttct 807  
|||||  
Db 424 CTGAAAGAGATGGTCCCAATGGCTCTGACCCACTGTGCTTACTATCTATATCTTTCT 483  
QY 808 catgtgacctggttaaaagacttgaattcagagcctcatctgagccctactagatagtaga 867  
|||||  
Db 484 CATGTGACCTGGTAAAGACTTGAATTCAAGGCTCATTTGAGGCCCTACTAGTATGAGA 543  
QY 868 gaaggagctctggccaaggaagaacacagacttgcacaaatttatacacttttggct 927  
|||||  
Db 544 GAAGGGAGTGTGCCAAGGAAGACACACACCTTGACAAATTTATATACCTTTTGTCT 603  
QY 928 gtaattatgaaggaaagtctggcactcagaaacaaagaaactcctgtatgcagagtagg 987  
|||||  
Db 604 GATATTTCATGAAGGAAATTTGGCACTCAGAAACAAAGAACTCTTGATGCGAGATAGG 663  
QY 988 gatgtgcatctgctcgggcctggtcctaaatgcacacagtcacatggttlatgtaaacagg 1047  
|||||  
Db 664 GATGCTCATCTCTCGGGCTGGCCCTAAATGACACAGTCAATGATGTTATGTAAGAGG 723  
QY 1048 tctctg---ccaaagtcgtatgtatgtccacaggaatacattatgtgcatgtgga 1104  
|||||  
Db 724 TCTCTGCTACCAAGGTCTGATTGGATGTCACAGGAAATCACTTATTTGCGATGATTGGA 783  
QY 1105 atggagccacccctggaagtgcacataatctcctgaaggtcacacatttctgtgag 1164  
|||||  
Db 784 ATGGGACCACTCTCTGAAGTGCATCAATTTCTCTGAAGGTACACATTTCTTTGAGG 843  
QY 1165 aacacatgcagagcgtctctgtgaatctgcacataaacttcccttactgtcacaacatc 1224  
|||||  
Db 844 AACCATGCGCAGCGTCTTGGAAATCTCGCAATATCTTCTTACTGCTCAACACTC 903  
QY 1225 ttgatgaccttggacagttctactggtttgtcatalactcttcccaacaaatgatggc 1284  
|||||  
Db 904 TTGATGAGACCTTGAGACGTTTCTACTGTTTGTGATATCTCTTCCACCAACATGATGGC 963

QY 1285 atggaagctatgtcaaaagtagacagctgtcccaaggaaaccccaactcaagatgaaat 1344  
|||||  
Db 964 ATGGAAGCTTATGTCAAGATGAGACAGCTGTCCAGAGAACCCCACTACGATGAATAAT 1023  
QY 1345 aatgaagaagcggaaagactatgatgatcttactgattcttgaaatgtatgttgcagg 1404  
|||||  
Db 1024 AATGAAGACGGGAAGACTATGATGATGATCTTACTGATCTGAAATGATGTGTGCAGG 1083  
QY 1405 ttgatgatgaacacttcccttctcttctatccaattgtctcagttgtccaagaagcacct 1464  
|||||  
Db 1084 TTTGATATGACACCTCTCTCTTATCCAAATTTGCTCAGTTGCCAAGAAACATCTCT 1143  
QY 1465 aaaaactggttacatactacatgtcgtcgaagaaggagagactatgctccctatgctc 1524  
|||||  
Db 1144 AAAACTTGGGACTTATCATTTGCTGTGAAGAGAGAGACTGGAGCTATGCTCCTTATGTC 1203  
QY 1525 ctgcgcccgatgacagaagttataaagttcaatalttgaaacaaatggccctcagcgatt 1584  
|||||  
Db 1204 CTGCCCCGATGCACAGACTTATAAAGTCAATATTTGAACATGAGCCCTCAGCGATT 1263  
QY 1585 gttaggaagtacaaaagaagtcgcgattatgtgcatacagaatgaacacttgaagctcgt 1644  
|||||  
Db 1264 GTTAGGAAGTACAAAAGATCCGATTTATGGCATACACAGATGAACCTTTAAGACTCGT 1323  
QY 1645 gaagctatcagcatgaatcagaagatcttggagccttacttcttggaggagtggagac 1704  
|||||  
Db 1324 GAAGCTATTTACGCTGATGATGAGATCTTGGACCTTTACTTTATGGGGAGATTGGAGAC 1383  
QY 1705 acaactgttgatlatalttaagaaatcaagcaagcagacataatacattaccctcaagya 1764  
|||||  
Db 1384 ACACGTTGATTTATTTAAGATCAAGCAAGACGACCATATACATCTACCCGACAGGA 1443  
QY 1765 atcaagtagtccgtctcttctgtatcaaggagatccaaaaggtgtaaaacatttgaag 1824  
|||||  
Db 1444 ATCAGTATGTCCGCTCTTGTATTAAGAGAGATTACAAAAGGTGTAAACATTTGAAG 1503  
QY 1825 gattttccaattcgtccagagaagaataatcaaatataatggaagagctgtagaagat 1884  
|||||  
Db 1504 GATTTTCCATTCTGCCAGAGAAATTTCAATATATAATGGACAGTGTATGAAGAT 1563  
QY 1885 gggccaaactaaatcagaatcctcgtgctgaccgctactactagtttcgttaatatg 1944  
|||||  
Db 1564 GGGCCAACTAAATCAGATCTCGGTGGCTGACCGGCTATTACTGTATGTTAAATATG 1623  
QY 1945 gagaagatctagcttaaggaactatggccctctctctcctcgtctcgaagaatctgta 2004  
|||||  
Db 1624 GAGAGAGATCTAGCTTCAGAGACTATGAGCCCTCTCTCATCTGTCTAACAAGAACTGTGA 1683  
QY 2005 gatcaagaagaagaacacagataatgtcagacaagagaatgtcatccgttttctgtaattc 2064  
|||||  
Db 1684 GATCAAGAGAGAAACCGATATATGTCTGACAGAGAAATGTCAATCTCTGTTCTGTATTT 1743  
QY 2065 gatgagaaccgaagctgtgtactcctcaagaaatatacaacgcttctcccaatccagct 2124  
|||||  
Db 1744 GATGAGAACCGAAGCTGTGCTACACAGAGAAATATACAAAGCTTTTCCCAATCCAGCT 1803  
QY 2125 ggaagtgaagcttgaagatccaaggttccaagctccaactcaatgtcagacatcaatggc 2184  
|||||  
Db 1804 GGAAGTGAAGTTGAGATCCAGATTCAGAGCTTCCAAACATCATGTGCAACATCAAGTGC 1863  
QY 2185 tatgtttgatagttgtgacgtgtcagttgttcttgataaggtgtgacatgtgatact 2244  
|||||  
Db 1864 TATGTTTGTATAGTTTGATGCTGATGTTGTTGTTGATAGGTGCGATACGTGTACATT 1923  
QY 2245 ctgaagcatggaacagactgtacttcttctgtcttcttctcttgatatacacttcaaa 2304  
|||||  
Db 1924 CTAAGCATTTGAGACACAGACTGTGACTTCTTCTGTCTTCTTCTGGAATATACCTTCAAA 1983  
QY 2305 caaanaatgtctatagaagcaacatccactatcccatctcagaagaaactgtcttc 2364  
|||||  
Db 1984 CACAAATAGTCTATGAAGACACACATCCATATTCCTCATTTCTGAGAGAAATGCTTTC 2043  
QY 2365 atgtcagtgtaaaaccaaagttctatgtattcttggggtgcccacaactcgaagcttctggaac 2424



```

Db 204 ATGCGATGAAAAACCCAGGCTATGATTTGGGGTCCACAACTGACACTTTGGAAAC 2103
Oy 2425 aagagcctgacgccttactactaaagttcttagtctgacaaagacacggtgattatc 2484
Db 2104 AGAGGCAATGACCGGCTTACTAAGGTTTCTAGTTGTGACAAAGACACTGTCATTTATAC 2163
Oy 2485 gagacacgtatgaagatatttcagcacttctgctgagtaaaacatgccaatgaca 2544
Db 2164 GAGGACAGTTATGAAGATTTTCAAGCATCTTGCTGAGTAAACAAATGCCATTGACCA 2223
Oy 2545 agaagcttcccccgaattcaagacacctagcactagcagaagaatltatgcaacc 2604
Db 2224 AGA----- 2226
Oy 2605 ccacagctcttgaaacgccatcaacgggaaataactgctactactcttcagtcacaa 2664
Db 2227 -----GAAATAACTGCTACTCTTCAAGTCAATCA 2259
Oy 2665 gaggaatctgactatgatacatalcatcagctgaatgaagaagaatttgacatt 2724
Db 2260 GAGGAATTTGACTATGATGATACCATATCAGTTGAATGAAGAAGATTTTACATT 2319
Oy 2725 tatgataagatgaataatcaagcccccgacgcttcaaaagaacacgacactattt 2784
Db 2320 TATGATAGGATGAATAATCAAGGCCCGCAGCTTTCAAAAGAAACAGACACTATTTT 2379
Oy 2785 atgctcagctgagagagccttgatataatgagatgagtaactcccaatgcttaaga 2844
Db 2380 ATTGCTCAGTGGAGAGGCTCTGGGATTTATGAGATGATAGTACTCCCAATGTTCTAGA 2439
Oy 2845 aacagggcccaagctggcagctgctcccaatgcaagaagctgttctccagaatctact 2904
Db 2440 AACAGGCTCAGAGTGGCAGTGCCTCAGTTCAGAAAGTGTTCAGAAATTTACT 2499
Oy 2905 gatggcccttactagcccttataccgtggagaaactaaatgtgaactttggaactctg 2964
Db 2500 GATGGCTCTTACTACGCCCTTATACCGTGGAGAACTAAATGAACTTTGGGACTCTCG 2559
Oy 2965 gggccatataaagacagaagctgaaagataatcatgctaaacttcagaatcaggcc 3024
Db 2560 GGGGCATATATGAAGACGAGAGTGAAGATATATCATGTAATCTTCAAGAAATGAGCC 2619
Oy 3025 tctgccttacttctctatctactagccttattcttaagaagaatcagaaggcaaga 3084
Db 2620 TCTGCTCCCTATTCCTTATTTAGCCTTATTTCTTATGAGAAATCAGAGCAAGGA 2679
Oy 3085 gcaagaacctgaaaaaaacttgtaagcctaaatgaacaaacttacttttgaagaatg 3144
Db 2680 GCAGAACTTGAATAAACTTGTCAAGCTTATGAAACCAAACTTACTTTGGAAAGTG 2739
Oy 3145 caacatcatatgacccaactaaagatgagcttgactcaagccttggtctattctct 3204
Db 2740 CAACATATATGACACCCACAGTAAGATGATGATTGACTGCAAAGCCTGGGCTTATTTCT 2799
Oy 3205 gatgtgaacctgaaaaaagatgtaactcaagcctgatttgacccctctgctgctgcaac 3264
Db 2800 GATGTTGACCTGGAATAAGATGTGCACCTCAGGCTGATGGAACCCCTTCTGCTGCAC 2859
Oy 3265 acaaacacactgaacctgtctcagtgaggagaaagtacagtaacaggaatttgctgtt 3324
Db 2860 ACTAACACACTGAACCCCTGCTCATGGGAGACAGTACAGTAAGGAATTTGCTGTGTT 2919
Oy 3325 t-----tcacacatcttgaatgaaccaaagctgctactcatgtaaaatatagaaagaac 3381
Db 2920 TTGCTCAGCATCTTTGATGAGACCAAAAGCTGTAATCTCACTGAAATATGGAAGAAAC 2979
Oy 3382 tgcagaggtccctgcaatlatcagaatggaagatccaccatttlaaagagaattatgcttc 3441
Db 2980 TGCAGGCTCTCCCTCATATATCCAGATGGAAGATCCCACTTTTAAAGAAATTTATGCTTC 3039
Oy 3442 catgcaatcaatgctatcataatgataactactgctgaagtaatgctccagaatcaa 3501

```

---

```

Db 3040 CATGCATCAATGCTACATATAATGGATACACTACCTGGCTTAGTAATGCTCAGATCA 3099
Oy 3502 aggaatcgatglatctgctcaagcagtgagcaaatgaaacatccatctatctatctc 3561
Db 3100 AGGATTGATGATCTCTCTCAGCAGTGGCAGCATGAAACATCCATTTGATTCATTTC 3159
Oy 3562 agtgcacatgcttcaactgtacgaaaaaagaggagataaaatgacatgtaacatctc 3621
Db 3160 AGTGACATGCTGTCACATGTAAGAAAAAGAGATGTAATAAATGGACCTGTAACATCTC 3219
Oy 3622 tatcaggtgtttttgaagacagtggaagtgttaccatccaaagtgtgaatttggcggtg 3681
Db 3220 TATCCAGGTGTTTTGAGACAGTGGAAATGTTACCATTCCAAAGCTGGAATTTGGCGGGTG 3279
Oy 3682 gaatgcttatctggagacatclcatatgcttggaatggaacacacttctctgcttgaagc 3741
Db 3280 GAATGCCCTTATTTGGGAGACATCTACATCTCGGATGAGACACACTTTTCTGCTACAGC 3339
Oy 3742 aataagctcagactcccttgggaatggtctctggaacatlaagatatctcagattaca 3801
Db 3340 AATAAGTGTCAAGACTCCCTGGGAATGGCTCTGACACATTTAGAGATTTTCAGATTACA 3399
Oy 3802 gcttcaggaacataatgagcagtgggccccaaagcttggccagactctatctccgagta 3861
Db 3400 GCTTCAGACATATGAGACAGTGGGCCCAAGCTGGCCAACTTCATTTATTCGGATCA 3459
Oy 3862 atcaatgctcgtgagaccaaagagcccttctcttgaatcaagtgatgctgttggacca 3921
Db 3460 ATCAATGCTGTGAGACCAAGAGCCCTTTCTTGATCAAGTGGATGCTGTGGACCA 3519
Oy 3922 atgattatccagcatalcaagaacccaggtgtccggtcagaagttctccagccttaac 3981
Db 3520 ATGATTAATTCAGGCGATTAACCCAGGCTCCGCTCAGAAAGTCTCCAGCCTCTACATC 3579
Oy 3982 tctcagttatcatatgataatgacttgaatggaagaagtggagactatcagaagaat 4041
Db 3580 TCTCAGTTATCATATGATATATAGTCTTGATGGAAGAAAGTGGACACTTATCAGGAAT 3639
Oy 4042 tccacttgaaccttaatgctctctcttggcaatgltggaatcatcttggataaacaacat 4101
Db 3640 TCCACTGGAACCTTAATGCTGCTTTGCAATGATGATTCATCTGGATTAACACAT 3659
Oy 4102 attttaaacctccaattatgtctcogataatccggttggacccaactatataagatt 4161
Db 3700 ATTTTAACTCCCTCAATATGCTGATACATCCGTTTGCACCCAACTATTAATGACTT 3759
Oy 4162 cgaagcactcttcagatgagttgatggctgttaattaaatglttgcacatggcaatg 4221
Db 3760 CGCAGCACTCTTCGATGAGATTGAGGCTGTGATTTAAATGATTGCAGCATGCCATTG 3819
Oy 4222 ggaatggaagtaaaagcaatatacagatgacagataactgcttcatctacttaccat 4281
Db 3820 GGAATGAGAGTAAGAAAGCAATATCAGATGACAGATTAATCTGCTCATCTTACCAAT 3879
Oy 4282 atgtttgcacctgtgtctcttcaaaagctcgacttcaactccaaagggagatgattgc 4341
Db 3880 ATGTTTGCACCTGGTCTCTTCAAAAGCTGCACTTCACTCCAAAGGAGAGTAAGGCC 3939
Oy 4342 tggagacctcagtggaataatccaaagagtggtctgaagtggaactccagaagaatg 4401
Db 3940 TGGAGACTTCAGGTGAATATATCAAAAGATGCTGCAAGTGGACTTCAGAGACATG 3999
Oy 4402 aaagtacaggaatgaactactcaaggagtaaatctctgcttaccagcatglatgtaag 4461
Db 4000 AAAGTCAAGAGATTAATCACTACGAGATTAATAATCTGCTTACACAGATGATGTAAG 4059
Oy 4462 gagttcctcatctccagcagtcgaagatggaatcaatgaggtctcttcttcaagaatggc 4521
Db 4060 GAGTTCCTCATCTCCAGAGTCAAGATGGCCATCAGTGGACTCTCTTTTTCAGAAAGGC 4119
Oy 4522 aaagtaaaagtgttttcaagggaatcaagactcttcaaaccttggtggaactctcagac 4581
Db 4120 AAAGTAAGGTTTTTCAAGGAATCAAGACTCTTCAACCTGTGGTGAACCTCTAGAC 4179

```



|    |      |   |      |
|----|------|---|------|
| Db | 1621 | GAGAGGATCTACCTTCAGAGCTCATTTGGCCCTCTCTCATCTGCTACAAAGATCTGTA        | 1680 |
| Qy | 2005 | gatlcaaaaggagaaaccagataaattgltcagacaagaagaatgltcatcgttlttctgtatlt | 2064 |
| Db | 1681 | GATCAAGAGGAACACGATATATGTGACACAAGAGGAAATGTCATCGTTTTCGTATTT         | 1740 |
| Qy | 2065 | gatgagaaccgaagcttgtgtatcctccacagagataatacaagcttcttcccaatccagct    | 2124 |
| Db | 1741 | GATGAAACCGAAGCTGCTGCTCCACAGAGAAATATACACCTTTTCCCAATCCAGCT          | 1800 |
| Qy | 2125 | ggagtcgaacttgaagatccaagagttccaaagcctccaactcatatgacacagataatggc    | 2184 |
| Db | 1801 | GGAGTCAGCTTAGATTCAGAGTTCCAGAGCTTCCAAACATCATGACACAGCATCAATGGC      | 1860 |
| Qy | 2185 | tatgttcttgatagttctgcagttgtcagtttgttctgcatgagtggtgacatactgtgtacat  | 2244 |
| Db | 1861 | TAATGTTTTGATGATGTTCAGAGTTGTGAGTTTGTTCATGAGGTGGCATATCTGGTACTT      | 1920 |
| Qy | 2245 | ctaagcatctggagacacagactgacttcccttctgtcttcttctcttgatataccttcaaa    | 2304 |
| Db | 1921 | CTTAAGCATTTGGACACAGATGACTTCTCTTCTGCTTCTCTGATATACCTTCAAA           | 1980 |
| Qy | 2305 | cacaaatggtctatgaagaacacatccatcttccatctcaggaagaacgtgtctc           | 2364 |
| Db | 1981 | CACAAATGCTTATGAGACACACTCACCTACCTATCTCCATTTCTCAGGAGAAAGTGTCTTC     | 2040 |
| Qy | 2365 | atgtcatctgaaaacaccgaagtcatagtatctgtgggtgcacaactcagacttccgagac     | 2424 |
| Db | 2041 | ATGTGATGGAAACCCAGGTCTATGATCTGTGGGTGCCCAACCTCAGACTTTCGGAGAC        | 2100 |
| Qy | 2425 | agaagcatgaaccgctctactgaagtttctagtgtgtacaagaacactgltgatlatac       | 2484 |
| Db | 2101 | AGAGGCATGACCCCTTACTGAGAGTTTCTAGTTGTGACAAAGACACGTGGATTATTAC        | 2160 |
| Qy | 2485 | ggaggaaagltatgaagatalattcagcatcttctgtgtaaaacaatgtccatttgaacca     | 2544 |
| Db | 2161 | GAGGACAGTTTATGAATATATTTTCAGCATACTTGTGAGTAAAAACAATGTCATTGAACA      | 2220 |
| Qy | 2545 | agaagcttctccagaat-----  | 2562 |
| Db | 2221 | AGAGCTTCTCCAGAGATCTCTTGTGCGATPACCACATATGCTACTCAGATACCAAAA         | 2280 |
| Qy | 2563 | -----   | 2562 |
| Db | 2281 | GAAGATGGAATCCCAAGAGAAGTACACGAAAAAACAGTTTAAAGAAAAGATATACC          | 2340 |
| Qy | 2563 | -----   | 2562 |
| Db | 2341 | ATTTGTCCCTGAACGCTTGTGAAGCAATCATGCAATAGCAGCAATTAATGAGGACAA         | 2400 |
| Qy | 2563 | -----   | 2598 |
| Db | 2401 | AATTAAGCCCGAATTAAGATGCACCTGGGCAAGCAAGGTGAGACTGAAAGGCTTGCTCT       | 2460 |
| Qy | 2599 | ggcaccaccaaccgltcttgaacgcacatacaagggaataatactgltactactctcagtca    | 2658 |
| Db | 2461 | CAAAACCCACAGTCTTTGAAAGCCCATCAACGGGGAATAACTGCTACTACTTCAGTCA        | 2520 |
| Qy | 2659 | galcaagaagaaatctgactatgatlacacatatcaglttgaatlgaaagaagaattct       | 2718 |
| Db | 2521 | GATCAAGAGGAATTTACATATGATATACATATCACTTGAATGAAGGAAGGAAATTTT         | 2580 |
| Qy | 2719 | gacatttgaatgaagatgaanaatccaagagcccccgcagcttccaagaanaaacgcgac      | 2778 |
| Db | 2581 | GACATTTATGATGAGGATGAATAATCAAGAGCCCCCGCAGCTTTCAAAAGAAACGACAC       | 2640 |
| Qy | 2779 | tatttattgtctgcaatgagagagctctggatataatgagatgagtgagtgagtgagtgagtg   | 2838 |
| Db | 2641 | TATTTTATTTGCTGCAGTGGAGAGGCTCTGGGATTTATGGGATGAGTACTTCCCCATGTT      | 2700 |
| Qy | 2839 | ctaagaanaacaggtctcagagtggtgcagtgctccctcagltccaagaagtgttctccagaa   | 2898 |
| Db | 2701 | CTAACAACAGGGCTCAGAGTGGAGTGTCCCTCAGTTCAAGAAAGTGTTCCTCAGAA          | 2760 |
| Qy | 2899 | ttactgaatggcttcccttactaagcccttataccgttggagaactaaatgaacaatttggaa   | 2958 |
| Db | 2761 | TTTACTGATGGCTCCTTACTACGCCCTTATACCGTGGAGAACTAATGACATTTGGGA         | 2820 |
| Qy | 2959 | ctctcggggccatatataaagacagaaatltgaagataataacatggttaacttcaagaat     | 3018 |
| Db | 2821 | CTCTGGGGCCATATATATAGAGCAAACTTTAAGATATATATCATGATCATTTTCAGAAAT      | 2880 |
| Qy | 3019 | caagcctctgtcccttacttcccttacttcccttacttcccttacttcttgaagaagacagag   | 3078 |
| Db | 2881 | CAGGCTCTGCTCCCTATTTCTCTTATTTCTGACCTTATTTCTTATGAGGAAGATCAAGC       | 2940 |
| Qy | 3079 | caaggaagcagaacctagaanaaactltgtcaagccataatgaaccaaacttactttgg       | 3138 |
| Db | 2941 | CAAGGACGAGAACTGAGAAAAACTTTGTCAAGCCTATATGAACCAAAACTTACTTTTGG       | 3000 |
| Qy | 3139 | aaagtgtcaacatcataatggtcacccaactaagaatgaatttgaactgtcaaaagcttggctat | 3198 |
| Db | 3001 | AAAGTGCAACATCATATGCGACCCCACTAAAGTGTGTGACTCMAAGCTGGGCTTAT          | 3060 |
| Qy | 3199 | ttctcttgatgttgaaccttggaaaagaatgtgaactcaagccgtatgtgaacccctctgtgc   | 3258 |
| Db | 3061 | TTTCTGTGATGTTGACCTGGAAAAAGATGTGCATCAGCCCTGATTTGGACCCCTTGCTC       | 3120 |
| Qy | 3259 | tgccacacataacacacttgaacccttgcataatgtggagacagtgaaagttcagaaatltgtc  | 3318 |
| Db | 3121 | TGCCACACTTAACACACTGACCCGTCTCATGSGAGACAAAGTACAGATACAGAAATTTGCT     | 3180 |
| Qy | 3319 | ctgtttt---tcaacatcttgaatgaagcaaaagcttvgtaacttcaactgaataatgtgaa    | 3375 |
| Db | 3181 | CTGTTTTCCCTCACCATCTTTGATGAGACCAAAAGCTGTACTTCCACTGAAAAATGTGAA      | 3240 |
| Qy | 3376 | agaaactgcaaggtctcccttgtaataataccagatlgaaagatcccaactttaaagaatat    | 3435 |
| Db | 3241 | AGAACTGCAAGGCTCCCTCGAATATTCAGATGGAAGTATCCACTTTTAAAGAAATATAT       | 3300 |
| Qy | 3436 | cgcttcaatgaatcaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaat    | 3495 |
| Db | 3301 | CGCTTCCATGCAATCAATGAGGCTTACATATGATGATACACTGCTGCTTATGATGCTCAG      | 3360 |
| Qy | 3496 | galcaaaagatcgaatgtgtatctgtcagcaatgtggacgaatgaanaacatcattctat      | 3555 |
| Db | 3361 | GATCAAAAGATTCGATGATCTGCTGCTGAGTGGGACACATGAAACATTCATTTATT          | 3420 |
| Qy | 3556 | catttcaatgtgaatcgtgttcaatcgttgaagaaaaaagagaggtataaaatgtgcctgtac   | 3615 |
| Db | 3421 | CATTTCACTGACATGTTGTCTACGTACGAAAAAAGAGGTATTAATAATGTGCACCTTAC       | 3480 |
| Qy | 3616 | aatcctcatccaggtgttcttgaagacgttgaataatgtlataccatccaaagcttgaatttg   | 3675 |
| Db | 3481 | AATCTCTATCCAGGTGTTTTTGAAGACATGTGAATGTTACATTCMAAGCTGGAATTTGG       | 3540 |
| Qy | 3676 | cggttgaagaatgcttatttgcagacatcactatgcttgcagatgaagacacttcttctgtg    | 3735 |
| Db | 3541 | CGGGTGAATGCTTATTTGCGGACATCTACATGCTGTGAGACACACTTTTCTGCTG           | 3600 |
| Qy | 3736 | tacagcaataagtgatgagactcccttggaaatgtcttcttgc                       |      |

QY 3976 tacatctcaagttatcatcatgtatagtttgatggagaagtggcagactatcga 4035  
|||||  
Db 3841 TACATCTCTCAAGTTATCATCATGTATAGTCTTGATGGAGAAAGTGGCAGACTTATCGA 3900  
QY 4036 ggaatctcactggaaacttaattgctctcttggcaattgtaatactctggataaaa 4095  
|||||  
Db 3901 GGAATATTCACCTGGAACTTAATGCTCTCTTGGCAATGGATGATCATCTGGATATAA 3960  
QY 4096 cacaatattttaacctccaattatgtctgcgataacatccgtttgacccaactaatat 4155  
|||||  
Db 3961 CACAATATTTTAACTCCCAATTATTGCTGATACATCCGTTTGGCACCACCACTTAT 4020  
QY 4156 agcatctgcagcactcttcgcagtgatgtatggcgtgtatataatgtttgcaagatg 4215  
|||||  
Db 4021 AGCATTTGCGACGACTCTTCCCATGGAGTTGATGGGCTGTATTTAATAGTTGCGAGATG 4080  
QY 4216 ccatltyggaattgagatlaaagaacatatacagatgacagataactgcttcatctt 4275  
|||||  
Db 4081 CCAATGGGAATGAGAGATTAAGCAATATGACATGACAGATTACTGCTCATCTACTTT 4140  
QY 4276 accaataatgttgcacactggtctctcttcaaaagctgcagcttcaactccaaggagagat 4335  
|||||  
Db 4141 ACCAATATGTTTGCACCTGCTCTCTTCAAAAGCTGACCTTCACTCAAGGAGAGAT 4200  
QY 4336 aatgctcgagagactcagtggaataatccaaaagatggtctgaagtggagcttccagaag 4395  
|||||  
Db 4201 AATCCCTGGAGACCTCAGGGTAATATCCAAAAGAGTGGCTGCAAGTGACCTTCCAGAA 4260  
QY 4396 acaatgaataacacagagtaactactcaagagataaactctctgtcttaccagatgat 4455  
|||||  
Db 4261 ACATGAATATCAAGAGATTAAGTACTACAGAGATAAATCTGCTTACACATGTAT 4320  
QY 4456 gtgaagagatctctcatctccacagatcaaatggtccatagtggactctcttcttctcag 4515  
|||||  
Db 4321 GTCAAGAGATTCCTCATCTCCACAGCAAGATGAGCCATGAGTGACCTCTTTTTCAG 4380  
QY 4516 aatggcaaatlaaagtttctcagggaaatcaagactcttcaactgtggtggaactct 4575  
|||||  
Db 4381 AATGGCAATTAAGGTTTTTTCAGGGAAATCAAGACTCTTTCACACCTGTGGTGAACCTCT 4440  
QY 4576 ctgaaccacagcttactgactcgtactcctcgaattcaaccacagatltggtgcacag 4635  
|||||  
Db 4441 CTGACCCACAGCTTACTGACTCTGCTACTCTGCAATTCACCCCAAGATTGGTGGACAC 4500  
QY 4636 attgacctgagatgagatgtctcgtgctcgagagcagagactcttaac 4683  
|||||  
Db 4501 ATTGCCCTGAGATGAGGTTCTGGGCTGCGAGGACAGAGACTCTAC 4548

RESULT 6  
LOCUS 108641 4551 bp PAT 02-DEC-1994  
DEFINITION Sequence 1 from Patent WO 8800831.  
ACCESSION 108641  
VERSION 108641.1 GI:588647  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4551)  
AUTHORS Pasetk,M.P.  
TITLE DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE  
JOURNAL POLYPEPTIDES IN HIGH YIELDS  
FEATUES Patent: WO 8800831-A 1 11-FEB-1988;  
location/Qualifiers  
source 1..4551  
BASE COUNT 1345 a 1003 c 1004 g 1198 t 1 others  
ORIGIN

Query Match 85.7%; Score 4085; DB 10; Length 4551;  
Best Local Similarity 95.1%; Pred. No. 0;  
Matches 4325; Conservative 0; Mismatches 31; Indels 192; Gaps 3;  
QY 328 gccacagaataactactcctggatgacgtggaactgtcatggaactatgcaaatgtat 387  
|||||  
Db 4 GCCACCGAAGATATCACTCTGGGTGCGAGTGAACCTGCATGGAGATATATGCAAGAT 63  
QY 388 ctggtgagctgctctggtggcgaagaattccctccctgaagtgccaaatttctcatc 447  
|||||  
Db 64 CTCGGTGAGCTGCTGTGGACGCAAGATTTCTCTGAGATGGCCAAATCTTTTCCATTC 123  
QY 448 aacacccatcgtgtgtaacaaagactctgttgtgaatctcaggttcccttcaac 507  
|||||  
Db 124 AACACCTCATCTGTTACAAAAGACTCTGTTGTACAAATTCACAGGATCACCTTTTAC 183  
QY 508 atcgtaagccaaagccacccttgatggtgctgtaagtgcttaccatccacagctgaagt 567  
|||||  
Db 184 ATGCGTAAAGCCAGGCGCACCTGGATGGTCTGTAGTCTTACATCCAGGCTGAGGTT 243  
QY 568 tatgatacagtggtcatataacttaagaacatggtctccatctcgtctcatgtct 627  
|||||  
Db 244 TATGATACAGTGGTCAATTAACCTTAAGAACATGGCTTCCATCCTGTCAAGTTCATGCT 303  
QY 628 gttggtatctctacttggaaagcttctgagggagctgaatatgattgtaagaccagca 687  
|||||  
Db 304 GTTGGTATCTCTACTGGAAGGCTTCTGAGGAGCTGAATATGATGATGACACCAAGTCAA 363  
QY 688 agggagaagaatgataaagctctccctggtggaagccatacatatgctgcaagtc 747  
|||||  
Db 364 AGGAGAGAAAGATGATTAAGTCTTCCCTGTTGGAAGCCATCATATGCTGGCAGGTC 423  
QY 748 ctgaagaagaatggtccaatggtccttgaccacatggtcttactcatatcttct 807  
|||||  
Db 424 CTGAAGAGATGGTCCAAAGGCTCTGACCCACTGCTTACTCATATCTTCTTCT 483  
QY 808 catggaacctgttlaaagacttgaattcagggcttcaattgagccctcagatgata 867  
|||||  
Db 484 CATGTGACCTGTTAAAGACTTGAATTCAGGCTCTATTGGAGCCCTACTAGTATGAGA 543  
QY 868 gaaggagcttggccaaagaaagacagacttgcacaaattataactacttctgt 927  
|||||  
Db 544 GAAGGAGTGTGGCCCAAGGAAAGACACACACTTGCAAAATTTATATCTTTTGGCT 603  
QY 928 gtaattgataaaggaagaattggtcactcagaacaaagaacctctgtagcaggaatg 987  
|||||  
Db 604 GTATTTGATGAAGGAAAGATTTGGCACTGCAAAACAAAGATCTTGATGCGAGATAGG 663  
QY 988 gatgtgcatctgctcgggcttgccctaaatgcacacagtcgaatgttattgtaaacag 1047  
|||||  
Db 664 GATGTGCAATCTGCTCGGGCTTGCCCTTAAATGACACACATGATGTTATGTAACAGG 723  
QY 1048 tctctg----cgaagtctgattgtaatgcacagaaatcaatctatgtgcattgtatgga 1104  
|||||  
Db 724 TCTCTGTACACAGGTCTGATTTGATGTCACAGGAATCACTATATTGGCATGATGATGGA 783  
QY 1105 atgggacacactcctgaatgtaactaataatctcctgaaggtacacacatttcttgag 1164  
|||||  
Db 784 ATGGGACACCACTCTCTGAAGTGCACATATTTCTCTGCAAGTGCACACATTTCTTGAGG 843  
QY 1165 aaccatcgacagcgctccttggaatctgcgcaataacttcttactgttcaaacctc 1224  
|||||  
Db 844 AACCATCGCAGGCGTCTTGGAAATCTGCCAATATCTTCTTACTGCTCAAAACATC 903  
QY 1225 ttgatgaccttggacagtttctactgtttgttcatalactcttccacacacatgtagtgc 1284  
|||||  
Db 904 TTGATGACCTTGACAGATTTCTACTGTTTGTGATATCTCTTCCACCAACATGATGGC 963  
QY 1285 atggaagcttatgtcaaaagtagacagctgtccagaggaacccaactacagaaatga 1344  
|||||  
Db 964 ATGGAAGCTTATGTCAAAAGTAGACAGCTGTCCAGAGGAACCCCACTTACCAATGAATAAT 1023  
QY 1345 aatgaagaagcggaaagactatgatagtatcttactgattctgaatgtatgtgtcagag 1404

```
|||||
Db 1024 AATGAAGAGGGAAGACATGATGATGATCTACTGATTTGGAATGATGTGCTCAGG 1083
Oy 1405 ttgatgatgacaacatcccttccttccttcaacatctgcagttgccaagaagatcct 1464
Db 1084 TTTATGATGACAACTCTCTCTCTCTTATCCAAATTCGCTGCTGCTGCTGCTGCTGCT 1143
Oy 1465 aaacttggatgacatctacatctgcttgaagaagagagactggagactgctccctta 1524
Db 1144 AAAACTGGGATACATTACATTTGCTGCTGAGAGAGACTGGAGCTATGCTCTCTTAGTC 1203
Oy 1525 ctgcccccgatgacagaagtataaagaatcaatactgacaatgcccctacggat 1584
Db 1204 CTCGCCCCCGATGACAGAGATTAAAGTCATTTGGAACAAAGGCCCTCAGCGGATT 1263
Oy 1585 gttggaagatcaaaaaagtcgatttatggcatcacagatgaaacctttaagaatcgt 1644
Db 1264 GGTAGGAGTACAAAAAAGTCGATTTATGGCATACACAGATGAACCTTTAGACTGCT 1323
Oy 1645 gaagctatcagcatgaaatcagaatcttggaccttacttacttctgaggagttgagac 1704
Db 1324 GAACGATTTGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1383
Oy 1705 acaatgttgattatattaaagaatcaagaagcagaacacatatacctacccctcagga 1764
Db 1384 ACACGTTGATTATATTAATAAATCAACCAAGCAGACCATATACATCTACCCCTACGGA 1443
Oy 1765 atcaatgatgctgcttctgttatcaagagagattaccaaaagtggtaaacatttgag 1824
Db 1444 ATCAGCTATGCTCCCTCTCTTATTTCAAGAGATTTACCAAAAAGTGTTAAACATTGGA 1503
Oy 1825 gatttccaatcttgcagagagaatataccaataataatgagcagtgactgtagaagat 1884
Db 1504 GATTTTCCAATTTCTGCCAGAGAAATATTTCAAAATTAATGAGCAGTACGCTAGAGAT 1563
Oy 1885 gggccaaactaaatcaagatccctcggtgcttgcctgacccgctatctactcagttcgtaatatg 1944
Db 1564 GGGCAACTAAATCAGATCCCTGGGCTGACCCGCTATCTGATGATGATGATGATGATG 1623
Oy 1945 gagaagatctagcttcaagagatcattggccctccctcactgctgataaagaatctgta 2004
Db 1624 GAGAGATCTAGCTTCAGGACTCATTTGGCCCTCTCTCATCTGCTACAAAAGATCTGTA 1683
Oy 2005 gatcaaaagagaaacagataatgtcagacaagagaaatgcatcctgttcttctgatt 2064
Db 1684 GATCAAGAGAGAAACCAAGATATGTCACAGAGAGATGTCATCTGTTTCTGTATTT 1743
Oy 2065 gatgagaacggaagctggttacctcaagagaataacaagcttctccccaatccagct 2124
Db 1744 GATGAGAACCGAAGCTGTACTCTACAGAGAAATATCAACGCTTTCCTCCCAATCCAGCT 1803
Oy 2125 ggaatgagcttggagatccagagatcccaagctcccaacatcagcagcatcaatggc 2184
Db 1804 GGAATGAGCTTGAGATCCAGAGATTCAGAGCTTCACCAATCATATGACAGCATCAATGGC 1863
Oy 2185 tatgttttgatagtttgcagttgtcagttgttctgcatgaggtggtcactggtacat 2244
Db 1864 TATGTTTGTATAGTTTGTGATGTCATGTTGTTTGTGATGAGTGGCATACGTAATTT 1923
Oy 2245 ctaagacttggagacagactgacttctctgttcttcttctctctgataccttcaaa 2304
Db 1924 CTAGAGCTTGGAGACAGACTGACTTCTTCTGCTTCTCTCTGATATACCTTTGAAA 1983
Oy 2305 caaaaaatgtctcatgagacacacacacacacacacacacacacacacacacacacac 2364
Db 1984 CAAAAATGCTTATGAGACACACTCACCTATTTCCATTTCTCAGGAGAAATCTTCTTC 2043
Oy 2365 atgtcgaatgaaaaaaccaagttctatgtatctgggtgtgccaacacacagacttccgaa 2424
Db 2044 ATGTCATGAGAAAAACCAAGGCTATGATGTTCTGGGGTCCCAACTCAGACTTTTGGAG 2103
Oy 2425 agaggaatgacgccttaactgaagttcttagttgtgacaagaacacgtgtgattatcac 2484
|||||

|||||
Db 2104 AGAGCATGACCGCTTACTGAGAGGTTTCTAGTTGTGACAAAGAACTGCTGATTTATAC 2163
Oy 2485 gaggaagttatgaagatatttcaagcacttgcgtgaataaacaatgcatgaaacca 2544
Db 2164 GAGGACAGTTATGAAGATATTTTACGATACACTTTCGATGATTAACCAATGCCATTGAACA 2223
Oy 2545 agaagctctcccaaat----- 2562
Db 2224 AGAAGCTTCCACAGATCTCTTCTGATTAACCACTATGTGACTCAGATACAAA 2283
Oy 2563 ----- 2562
Db 2284 GAAGGTGGAATTCOAAGAGAGTCAACGAGAAAAACAGCTTTTAAAGAAAAAGATACC 2343
Oy 2563 ----- 2562
Db 2344 ATTTTGTCCCTGACGCTTGTGAAGCATATGTCATAGCAGATTAATGAGGACAA 2403
Oy 2563 -----tcaagacacctagcactaggcaaaagcaattaat 2598
Db 2404 AATAAGCCGGAATAGAGTCACTGGGCAAGCAAGTAGACTGAAGGCTGTGCTCT 2463
Oy 2599 gcaaccccaagctcttgaacgcgcatacgaaggaataactgttactacttctcagta 2658
Db 2464 CAAAACCCACAGTCTTGAAAGCCATCAACGGAAATTAACCTGCTACTCTTCACTCA 2523
Oy 2659 gatcaagagaatgtactgatatcaatcactgattgaatgaaagaagaatctt 2718
Db 2524 GATCAAGAGGAATTTGACTATGATGATGATGATGATGATGATGATGATGATGATGATG 2583
Oy 2719 gacatttatgtagatgaaatcaagaagcccgagcttcaaaagaagaacagacac 2778
Db 2584 GACATTTATGATGAGAGATGAAGAAATCAGAGCCCGAGCTTTCAAAAGAAACACGACAC 2643
Oy 2779 tattttatgtctgaagtgagaagctctggatattgagatgagatgagatcccaatggt 2838
Db 2644 TATTTTATTTGCTGAGTGAAGAGGCTCTGGGATTTAGGATGATGATGATGATGATG 2703
Oy 2839 ctaagaacaaggtctcagatgagtgctccctcagttcagaagaagtgttctccaagaa 2898
Db 2704 CTAAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTGTGTTTCCAGGA 2763
Oy 2899 ttactatggtctccttactcagcccttataccgttggagaactaaatgaacatttggga 2958
Db 2764 TTTACTGATGCTCTCTTACTACAGCCCTTATACCGTGAGAACTAAATGAAACTTTGGGA 2823
Oy 2959 ctccgtggccatataataagagcagaagtgtgaataatcatcagtaacttccaagaat 3018
Db 2824 CTCTGGGGCCATATATTAAGAGCAGAAAGTTGAAGATATATCATGATTTTCAAGAAAT 2883
Oy 3019 caggtctctgctccatctccttctatctagccttattcttctatgagaagatcaag 3078
Db 2884 CAGGCTCTGCTCCCTATCTCTTATTTCTAGCTTATTTCTTATGAGAAAGTACAGAG 2943
Oy 3079 caaggaagcagaacctagaaaaaacttgtcgaagcttaatgaacaaacttacttgg 3138
Db 2944 CAAGGACGAGAACTAGAAAAAATTGTCTCAAGCCATATGAACCAAAACTTCTTTGG 3003
Oy 3139 aaagtgcacacatcatatagaccacacacacacacacacacacacacacacacacacac 3198
Db 3004 AAAAGTGCACATCATATGAGCACCACCTAAGATATGATTTGATGCAAAAGCTTGAGCTTAT 3063
Oy 3199 ttctctgaatgttgaacagaaaaagatgtgacacgaagcctgatttgaacccctctgtc 3258
Db 3064 TTCTCTGATGTTGACCTTGGAAGAAAGATGTGACACTCAGGCTGATTTGGAACCCCTTGTGTC 3123
Oy 3259 tgcacacttaacacactgaacctgtctcatgtggagacaagtgtgacgtacagaatctgtc 3318
Db 3124 TGCCACACTTAACAACATGAAACCTGTCTCATGTGAGACAAAGTACAGATGAAATTTGCT 3183
Oy 3319 ctgtttt---tcacacatttggatgagaccaaaagcgtgtacttctactgaaatatggaa 3375
Db 3184 CTGTTTTCCTACATCTTGTGATGAGACCAAAAGCTGGTATCTTCACTGAAATAATGGAA 3243
|||||
```

|    |      |   |      |
|----|------|---|------|
| OY | 3376 | agaactgcagggctccctgcgaatccgaatgysaagatccactttaaagaatlat           | 3433 |
| Db | 3244 | AGAAATCGCAGGGGCTCCCTGCAATATCCAGATGGAAGATCCACTTTTAAAGAAATTTAT      | 3303 |
| OY | 3436 | cgctccatgcgaataatgctcctataaagtgataactaccgtgttagatgagccag          | 3495 |
| Db | 3304 | CGCTTCCATGCATTCANATGGCTTACATAAAGATACACTACCTGGCTTAGATGATGANTCAG    | 3363 |
| OY | 3496 | gataaaagattcgaatgtaatactgcctcagacatgaggcaagaatganaaatccatctcat    | 3555 |
| Db | 3364 | GATCAAAAGAAATTCGATGATGTAATCGCTCAGACATGGGAGCAATGAAGAAATCATTTATTTAT | 3423 |
| OY | 3556 | catttcagtgagacatgctgtctcaactgtaagaaaaaagagagatataaaatgcaatgctgtac | 3615 |
| Db | 3424 | CATTTCAGTGGCAATGCTTCACTGTACCAAAAAAGAGAGATATAAATGACACTGTAC         | 3483 |
| OY | 3616 | aactccatccagtgcttcttgagaacagtgysaaatgtaaccaaccaagctcgysaatcttg    | 3675 |
| Db | 3484 | AATCTCATTCAGAGTGTGTTTTGAGACAGTGAAGATCTTACATCCAAAGCTGGAATTTGG      | 3543 |
| OY | 3676 | cgaggtagatgcctatcttgcgagacatcacatgctggatgtagacacatttctgtg         | 3735 |
| Db | 3544 | CGGCTGGAATGCCCTTATTTGGCAGCATCTACATCGCTGGATGACACACTTTCTGGTG        | 3603 |
| OY | 3736 | tacagcaataagtgtaagactccccctggysaatgyscttctgysaacatttagagattctag   | 3795 |
| Db | 3604 | TACAGCAATTAAGTGTACAGACTCCCCCTGGGAATGGCTCTGGACACATTAAGATTTTCAG     | 3663 |
| OY | 3796 | attaagcttcagagacaataatgagacagtgaggcccaaaagcttggccagactcatatcc     | 3855 |
| Db | 3664 | ATTACACTCTTCAGAGCAATATGTGAGCAGTGGGCCCAAGACTGGCAGACTTCATTATTTC     | 3723 |
| OY | 3856 | ggatccaataatgctcctggysagcaaccaagggcccttctcttgatcaagtgatctgtg      | 3915 |
| Db | 3724 | GGATCAATCATGCTCTGGAGACACCAAGAGGCCCTTTCTTGATCAAGTGTGATCTGTG        | 3783 |
| OY | 3916 | gcacaaatgattatcaagcgcacaaagaccagagtgcccgatcaagattctcagactc        | 3975 |
| Db | 3784 | GCACCAATGATTTATTCACGGCAGTACAGACCAGAGGTCGCCGTGCACAAATTTCCAGGCTC    | 3843 |
| OY | 3976 | tacatctcagttatcatcatcatgataagtlcttgatbgysaagaatgtagaactatcga      | 4035 |
| Db | 3844 | TACATCTCTAGATTTATCATCATGATTAATCTGTGATGGGAAGATGCGACACTATTACGA      | 3903 |
| OY | 4036 | gysaatccactggaaccttaatgysctctctcttgysaatgysgatctgtaactgysgaataaa  | 4095 |
| Db | 3904 | GGAATTTCCACTGGAACTTAATGGTCTTCTTTGGCAATGTGATTCATCTGGGATTAAGA       | 3963 |
| OY | 4096 | cacaatatltttaacctcccaattatctgcatatcatcogtlttgacaccaaccatcat       | 4155 |
| Db | 3964 | CACAAATTTTAAACCTCCCAATTATTTGCTCATTCATCCATCCGTTTGCACCCAACTCATTTAT  | 4023 |
| OY | 4156 | agacttcgysagacttctgcagatgagatgtgatgggctgtgatttaaatagtltgaagatg    | 4215 |
| Db | 4024 | AGCAATTCGACGACTCTTCCCATGGAGTGTGAGGGCTGTGATTTAAATAGTTGACGATG       | 4083 |
| OY | 4216 | ccattgysgaatgysagtaaaagaacatcagaatgcaagatctactgcttaactctt         | 4275 |
| Db | 4084 | CCATTGGGAATGGAGATGAAGCAATATCAGATGCAAGATTACTGCTTACTCTT             | 4143 |
| OY | 4276 | accaatatgtttgcaactggtctccttcaaaagctcgactcaactccaaagysagat         | 4335 |
| Db | 4144 | ACCAATATGTTTTGCCACCTGCTGCTCCTTCAAAAAGCTCGCACTTCACCTCCAAAGGAGAGAT  | 4203 |
| OY | 4336 | aatgctcgysagactcaagtgysaataaccaaaagtgysctgysaagtgysaattccagaag    | 4395 |
| Db | 4204 | AATTCCTGGAAGACTCAGTGGATTAATCCAAAAGATGGCTGCAAGTGGACTTTCCAAAG       | 4263 |
| OY | 4396 | acaaatgaagtcacagagatlaactactcagaggatlaaaatctctgcttaccagatglat     | 4455 |
| Db | 4264 | ACAAATGAATCTCAGAGATTAATCTACTGAGGAGTAAATCTCTTACAGAGATAT            | 4323 |

|                       |  |  |                |
|-----------------------|--|--|----------------|
| QY                    | 4456                                   | gtgaagaggttcctcatctcgaagtcgaagatgagtcacgtgacgtctctttttcaag         | 4515           |
| DB                    | 4324                                   | gtgaaggaggttcctcatctcgaagtcgaagatgagtcacgtgacgtctctttttttcag       | 4383           |
| QY                    | 4516                                   | atggcctaagtaaggttttttcagggaatacaagctcttctacaccgtgtgtgacgtc         | 4575           |
| DB                    | 4384                                   | aatggcctaaagtaaggttttttttcagggaatacaagctcttctacaccgtgtgtgacgtc     | 4443           |
| QY                    | 4576                                   | ctagaaccacgttactactgcgtacaccttgcgaattcaaaccccccaagtttgagtcacag     | 4635           |
| DB                    | 4444                                   | ctagaaccacgttactactgcgtacaccttgcgaattcaaaccccccaagtttgagtcacag     | 4503           |
| QY                    | 4636                                   | attgcacctgaagatggaagttcttggtgctgcgaggcacaggaacctctac               | 4683           |
| DB                    | 4504                                   | attggccctgagatggaagttcttggtgctgcgaggcacaggaacctctctac              | 4551           |
| RESULT                | 7                                      |  |                |
| AR034084              | AR034084                               | 5035 bp  | DNA            |
| LOCUS                 | AR034084                               | Sequence 1 from patent US 5869292.                                 | PAT            |
| DEFINITION            | AR034084                               |  |                |
| ACCESSION             | AR034084.1                             | GI:5949689   |                |
| VERSION               | AR034084.1                             | GI:5949689   |                |
| KEYWORDS              |  |  |                |
| SOURCE                | Unknown.                               |  |                |
| ORGANISM              | Unclassified.                          |  |                |
| REFERENCE             | 1 (bases 1 to 5035)                    |  |                |
| AUTHORS               | Voorberg,J.J.                          |  |                |
| TITLE                 | Hybrid proteins with modified activity |  |                |
| JOURNAL               | Patent: US 5869292-A 1 09-FEB-1999;    |  |                |
| FEATURES              | Location/Qualifiers                    |  |                |
| source                | 1..5035                                |  |                |
| BASE COUNT            | 1484 a                                 | 1127 c   | 1110 g         |
| ORIGIN                |  | 1314 t   |                |
| Query Match           | 78.7%;                                 | Score 3748;  | DB 9;          |
| Best Local Similarity | 87.4%;                                 | Pred. No. 0;   | Length 5035;   |
| Matches 4574:         | Conservative                           | 0;   | Mismatches 60; |
|                       |  |  | Indels 570;    |
|                       |  |  | Gaps 2;        |
| QY                    | 266                                    | ccacatgcaaatagagctctccacgtctctctctctgtgaccttttgatctgctta           | 325            |
| DB                    | 30                                     | ccacatgcaaatagagctctccacgtctctctctctgtgaccttttgatctgcttta          | 89             |
| QY                    | 326                                    | gtgcacacgaagaatactactacctgggtgcagtggaaccttcatbgygacatatabgaagt     | 385            |
| DB                    | 90                                     | gtgcacacgaagaatactactacctgggtgcagtggaaccttcatbgygacatatabgaagt     | 149            |
| QY                    | 386                                    | acctgggtgagctgtccttctgagcgaagattccctccataggtgcacaaatctttcat        | 445            |
| DB                    | 150                                    | acctgggtgagctgtccttctgagcgaagattccctccataggtgcacaaatcttttcAT       | 209            |
| QY                    | 446                                    | tcaaacacctcagctcgtgttacaacaaagacctctgttgtagaattcaaggttcaaccttca    | 505            |
| DB                    | 210                                    | tcaaacacctcagctcgtgttacaacaaagacctctgttgtagaattcaaggttcaaccttca    | 269            |
| QY                    | 506                                    | acatcgtctaagccaagccacaccttgcgatbgtctgtcgtagttcctaaccatcgaagctgag   | 565            |
| DB                    | 270                                    | ACATCGCTTAAGCCAGGCGCACCCCTGGATGGGCGTCTGAGTTCACATCCAGCGCTGAGG       | 329            |
| QY                    | 566                                    | tttatgatacagtggttcatcttaacttaagaacatbgtcttcccatctctgtcaagttctcat   | 625            |
| DB                    | 330                                    | TTTATGATTAACAATGGTGTCAATTACACTTAAGAACATGCGCTTCCATCTCTCAAGTCTTCAATG | 389            |
| QY                    | 626                                    | ctgttggtgtatccctaccggaagaccttcttgagggagctcgaaatagatgatcagaccagtc   | 685            |
| DB                    | 390                                    | CTGTGTGGTGTATCTCTACTGCGAAAGCTTTGTGAAGGAGCTGGAATATGATATATCAGACCAATC | 449            |
| QY                    | 686                                    | aaaggagaaagaaatgataaagttctccctcgtgtgagaagccatatatgtctgcacgg        | 745            |
| DB                    | 450                                    | AAAGGAGAAAGAAATGATGTTAAAGCTTTCCTCGGAGGGAAGCCATCATATGTTCTGGCAGG     | 509            |







Db 2670 TCCAAATTAGATTAAATGAGAAACGTGGGACAACGTGCAGATCCTTGGTTGGGATAACC 2729  
QY 2604 ----- 2603  
Db 2730 ACTATGCTACTCAGATACCAAAAAGAGTGTGAATCCCAAGAAAGTCACCAAAAAA 2789  
QY 2604 ----- 2603  
Db 2790 CAGCTTTAAGAAAAAGATACCATTTTGTCCCTGAACGCTTGTGAAGCATCATGCAA 2849  
QY 2604 ----- 2603  
Db 2850 TAGCAGCAATAAATGAGGACAATAATAGCCCAATAGAACTCACCTGGGCAACCAAG 2909  
QY 2604 ----- 2603  
Db 2910 GTAGGACTGAAAGGCTGTGCTCTCAAAAACCCACAGCTTTGAAACGCCATCAACGGGAAA 2969  
QY 2636 taactcgtaactcttcagtcagatcaagaaggaattgactatgatatcatatcaag 2965  
Db 2970 TAAGTGTACTACTCTTCACTGATCAAGAGAAATGACTATGATGATACCATATCAG 3029  
QY 2696 tgaagaatgaagaagaagatttgacattatgatagaatgaagaatcagaagcccgca 2755  
Db 3030 TTGAAATGAAAGAGAAATTTTGACATTATGATGAGATGAATAATCAGAGCCCGCA 3089  
QY 2756 gcttccaagaagaacacagacactatcttctgcagatgagaagctctgagatatag 2815  
Db 3090 GCTTTTAAAAAGAAACGACACTATTTATGCTCAGAGGAGAGCTCTGGGATTAATG 3149  
QY 2816 ggaatgaatgcctccacatgcttcttaagaagaagagctcagaatgagctgcctcagt 2875  
Db 3150 GGAATGAGTACTCTCCACATGTTCTTAAGAAACAGGCTCAGAGTGGCAGTGTCTCAGT 3209  
QY 2876 tcaagaagltgttctccaggaatttactgataagtgctccttactcagccctatacagltg 2935  
Db 3210 TCAAGAAAGTGTGTTCCAGGAATTAAGATGAGTGTCTTACTCAGCCCTATACCGTG 3269  
QY 2936 gagaagaatagacaatttggagctctcgtgggcacatatagaagacagaagttgaagata 2995  
Db 3270 GAAAGCTAAATGAAACATTTGGGACTCTCGGGGCATATATAAGAGAGAAAGTTGAAGATA 3329  
QY 2996 atatcatgtaacttccagaatcagaagcctcgtccctacttctcttcttactcagccta 3055  
Db 3330 ATATCATGTGAATTTTCAGAAATCAGGCCCTGCTGCCATTTCTCTATTTACGCTTA 3389  
QY 3056 ttctctatgagaagaatcagaagcagaagcagaacctagaanaaacttgtcaagccta 3115  
Db 3390 TTTCTTATGAGAGATTCAGAGGCAAGGACAGCAACCTAGAAAAAACTTTGTCAAGCTTA 3449  
QY 3116 atgaaccacaacttacttcttggaaagtgcaacatcatatgacacccactaaagaatgagt 3175  
Db 3450 ATGAAACCAAAACTTACTTTTGGAAAGTGCACATCATATGCGACCCCTAAAGATGAGT 3509  
QY 3176 ttgactgcgaagcctggtctatcttctgatagttgacctggaagaagaatgtgctactcag 3235  
Db 3510 TTGACTGCAAAAGCGCTGGCTTATTTCTGTGATTTGACCTGGAAAAAGATGTGCATCTCAG 3569  
QY 3236 gectgatggaaccctctggtctgcacacactaacaactgaaacccctgcatgagagagc 3295  
Db 3570 GCGCTGATTTGGACCCCTTCTGTGCTGCACACTTAACACTGAAACCTGTGCTCATGGAGAGC 3629  
QY 3296 aagtaacatagagaagattgtctgttttttcaacatcttltgaatgagcaaaagctgtgt 3355  
Db 3630 AAGTGAAGTACAGGAATTTGCTCTGTTTTCACATCTTTGATGAGACCCAAAGCTGGT 3689  
QY 3356 acctcaactgaaatatgagaagaagaactgcagagctccctgcaatatcaccagatggaatc 3415  
Db 3690 ACTTCACTGAAATATGAGAAAGAACTGCAGGCGCTCCCTGCATATTCAGAGTGAAGATC 3749  
QY 3416 ccaactttaaagaatcagcttccatgcaatcaaatggtctacataatgatatcaactac 3475  
|||||

Db 3750 CCACCTTTAAGAGAAATTAATCGCTTCACATCAATGAGTGTACATTAATGATACACTAC 3809  
QY 3476 ctggtctagtaatggtctcaggaatcagaagattcgatgfatctgctcacaatgaggaga 3535  
Db 3810 CTGGCTTAGTAATGGCTCAGGATCAAGATTCAGATGTATGCTCTCACATGTGGCAGA 3869  
QY 3536 atgaacaatccatctatctatctcagtgagacatgtgtcactgtacgaaaaaagaag 3595  
Db 3870 ATGAAACATCCATCTTATTTCACTTTCAGTGCAGATGTGTTCCATGTACGAAAAAAGAG 3929  
QY 3596 agataaataatgacatgtaacaatctctaccaggtgttcttgagacatggaatgttac 3655  
Db 3930 AGTATTAATAATGGCACTGTCAATCTCTATCCAGAGTGTGTTTGAGACAGTGAATAATGTAC 3989  
QY 3656 catcacaagcttgaaatttggcgggtggaatgctcttcttgaggagatctacatgctggga 3715  
Db 3990 CATCCAAAGCTGGAATTTGGCGGGTGAATGCTTATTTGGCGAGATCTACATGCTGGGA 4049  
QY 3716 tgaagcaactcttctggtgtacagaacaatagtgtaagactccctcgggaatggtctctg 3775  
Db 4050 TGAGCACTCTTTTCTGGGTACAGCAATTAAGTGTACAGACTCCCTGGGAATGCTTGTG 4109  
QY 3776 gacacatagagaatctcagatcacgcttcaggaacaatatgagacagtggcccaagc 3835  
Db 4110 GACACATTGAGATTTTTCAGATTAAGCTTCAGAGCAATAATGAGACAGTGGCCCCAAAGC 4169  
QY 3836 tggccagacttcaatctccggaatcaaatgacctggaggacccaagggcccttctct 3895  
Db 4170 TGGCCAGACTCTCAATTTTCGGATCAATCAATGCCGTGGAGCACAGAGGCCCTTTTCTT 4229  
QY 3896 ggaatgaagtgatctgtctggaaccaatgataatcaaggaatcaagaccaggtgtccc 3955  
Db 4230 GGAATCAAGTGGATGCTGTGTGGCACCAATGATTAATTCACGGCATCAAGACCGAGTGGCC 4289  
QY 3956 gtcaagaagltctcagcctctacatctcaglttalcatalgatalcttgatggga 4015  
Db 4290 GTGAGAAAGTGTTCACAGCCCTACATCTCAGTTTATCATCATGTAATCTTGATGGGA 4349  
QY 4016 agaagtggaagactatcgaagaatctcactggaaccttaatggtctcttggaaatg 4075  
Db 4350 AGAAGTGGAGACTTATTCAGAGGAATTCACATGGAACCTTAATGAGTCTTTGGCAATG 4409  
QY 4076 tgaatcactctggaataaacaacaatatlttlaacctccaattatgtctgatacatcc 4135  
Db 4410 TGGATTCATCTGGGATTAATAACAATATTTTAAACCTCCCAANTATTTGCTGCAATATCC 4469  
QY 4136 gtttcacccaactcaatcaatgcatcgcagaacctctgcagatgagttgatggctgtg 4195  
Db 4470 GTTTGCACCCCACTCATTTATGATTCGAGCACTCTTGCGATGGATGGCTGTG 4529  
QY 4196 atttaaatgttcagacatgacatgttggaatgagatgaagaatataatgataatgacaaga 4255  
Db 4530 ATTTAATATGTTGACACATGACATTTGGAAATGAGATTAAGCAATATCAGATGCAACGA 4589  
QY 4256 ttactgtctcatcacttcaacaataitgttggccacctggtctccttcaaaagctcgac 4315  
Db 4590 TTAAGTCTTCACTCATCTACTTACCAATATGTTTGGCCACGCTGCTCTTTAAAAAGCTCGAC 4649  
QY 4316 ttcacctccaagggagaglaatgcttgaagacctcagaatgaaatlaatccaagaagtgagc 4375  
Db 4650 TTTCACTCCCAAGGAGAGTAATGCTTGAGAGACCTCAGTGATTAATCCAAAAAGATGGGC 4709  
QY 4376 tgcgaatggaatctcagaagaacaatgaaagtcaacaggaaataactactcaaggagtaaat 4435  
Db 4710 TGAAGTGAACCTTCCAGAAAGACATGAATAAGTCAAGAGATTAATCTCAGGGAGTAATAT 4769  
QY 4436 cctcgtctaacagcatgatagtgaaagagttctcaatccacagcagtcgaagtgtgacac 4495  
Db 4770 CTTCTGTTACAGCAATGATGGAAGAGTTCCTCATCTCCAGCAAGTCAAAAGTGGCATC 4829  
QY 4496 agtgaactctcttcttccagaatggaagaataaggttcttcaagggaataatcaagactcct 4555  
Db 4830 AGTGAACCTCTTTTTCAGAAATGGCAAAAGTAAAGGTTTTCAGGGAATTAAGACTCTCT 4889  
|||||

|                       |  |   |                |
|-----------------------|--|---|----------------|
| QY                    | 4556                                   | taaccacgtctggtgtaactctcttagagcccaacgcttctctgctctgtaacctgtaattcaac | 4615           |
| Db                    | 4890                                   | ttcaacacctgtgttggaactctcttagaccacacgcttactgtaactgtaacctgtaattcaac | 4949           |
| QY                    | 4616                                   | cccagaatctgggtgcacacagatctgcaccttaaggaatggaatctctgagctgcgaagcagag | 4675           |
| Db                    | 4950                                   | cccagagttgggtgcacacagatctgcaccttaaggaatggaatctctgagctgcgaagcagag  | 5009           |
| QY                    | 4676                                   | acctctactgactgagagagttc   | 4699           |
| Db                    | 5010                                   | acctctactgagagagttc   | 5033           |
| RESULT                | 8                                      |   |                |
| AK071306              | AR071306                               | 5035 bp   | DNA            |
| LOCUS                 | Sequence 1                             | from patent US 5910481.   | PAT            |
| DEFINITION            | AR071306                               |   | 18-FEB-2000    |
| ACCESSION             | AR071306.1                             | GI:7222194  |                |
| VERSION               |  |   |                |
| KEYWORDS              |  |   |                |
| SOURCE                | Unknown.                               |   |                |
| ORGANISM              | Unknown.                               |   |                |
| REFERENCE             | Unclassified.                          |   |                |
| AUTHORS               | 1 (bases 1 to 5035)                    |   |                |
| TITLE                 | Voorberg, J.J.                         |   |                |
| JOURNAL               | Hybrid proteins with modified activity |   |                |
| FEATURES              | Patent: US 5910481-A 1 08-JUN-1999;    |   |                |
|                       | Location/Qualifiers                    |   |                |
|                       | 1..5035                                |   |                |
|                       | /organism="unknown"                    |   |                |
| BASE COUNT            | 1484 a                                 | 1127 c  | 1110 g         |
| ORIGIN                |  |   | 1314 t         |
| Query Match           | 78.7%;                                 | Score 3748;   | DB 9;          |
| Best Local Similarity | 87.4%;                                 | Pred. NO. 0;  | Length 5035;   |
| Matches 4374;         | Conservative                           | 0;  | Mismatches 60; |
|                       |  |   | Indels 570;    |
|                       |  |   | Gaps           |
| QY                    | 266                                    | ccacatcgtgaatataagagctctccacctgcttcttctgctcttcttgagatctgctta      | 325            |
| Db                    | 30                                     | ccacatcgtgaatataagagctctccacctgcttcttctgctcttcttgagatctgctta      | 89             |
| QY                    | 326                                    | gtgcacacagaaataactacccctgggtgcagtggaactgtcactgaggaactatgcaaatg    | 385            |
| Db                    | 90                                     | gtgcacacagaaatgactgctggtggcgactgtggaactgtcactgaggaactatgcaaatg    | 149            |
| QY                    | 386                                    | atctcgtgtagctgtgcctgtggaagcaaatctctctcctagagtgcccaaatctttccat     | 445            |
| Db                    | 150                                    | atctcgtgtagctgtgcctgtggaagcaaatctctctcctagagtgcccaaatctttccat     | 209            |
| QY                    | 446                                    | tcaaacacctcgaatcgtgttcataaaagacctcttgttagaatcaacogtctacatttca     | 505            |
| Db                    | 210                                    | tcaaacacctcgaatcgtgttcataaaagacctcttgttagaatcaacogtctacatttca     | 269            |
| QY                    | 506                                    | acatcgtcctaaagagagccacctgtagtgcctgtagtgcctacatccacacagctgag       | 565            |
| Db                    | 270                                    | acatcgtcctaaagagagccacctgtagtgcctgtagtgcctacatccacacagctgag       | 329            |
| QY                    | 566                                    | tttatgatacagtggtlcatltaactaaagacatggcttccatccctgtaagcttcaag       | 625            |
| Db                    | 330                                    | tttatgatacagtggtlcatltaactaaagacatggcttccatccctgtaagcttcaag       | 389            |
| QY                    | 626                                    | ctgttggtgtagtccctactggaagagctctcgaaggagctgaatatgatactgagcaagtc    | 685            |
| Db                    | 390                                    | ctgttggtgtagtccctactggaagagctctcgaaggagctgaatatgatactgagcaagtc    | 449            |
| QY                    | 686                                    | aaagggagaaagaaatgataaagtcttccctcgtgtgaaagcacaatacatgtctgcaag      | 745            |
| Db                    | 450                                    | aaagggagaaagaaatgataaagtcttccctcgtgtgaaagcacaatacatgtctgcaag      | 509            |
| QY                    | 746                                    | tccctgaaagaaatgataaagtcttccctcgtgtgaaagcacaatacatgtctgcaag        | 805            |

|    |      |   |      |
|----|------|---|------|
| Dp | 510  | TCCTGAAGAGAAATGTCATATGSCCTCTGACCCACAGTCGCTTACTACTACTATATCTTT      | 569  |
| Qy | 806  | ctcatctgacccttggtaaagaacttgaaatcaggccctcatctggagccctaagtatgta     | 865  |
| Dp | 570  | CTCATGTGTGACCTGGTTAAAGACTTGAAATTCAGGCCCTCATTTGGAGCCCTACATGATTA    | 629  |
| Qy | 866  | gagaaggaagctctggccaaggaagaagcacagaaccttgacaaattatacacttttg        | 925  |
| Dp | 630  | GAGAAAGGAGTCCTGCCCCAAGAAAGACACAGACCTTGCACAAAATTTATATCACTTTTGG     | 689  |
| Qy | 926  | ctgtaattgatagaaaggaagaagcttgcacccaaagaacaaagaacctctgatagcagata    | 985  |
| Dp | 690  | CTGATTTTGATGAAGGAGAAAGTTGGCAGCTCAGAACCAAGAACTCTTGATGACAGATA       | 749  |
| Qy | 986  | ggagatgctgtaclctcgtctcggcctgyccttaaaatgacacagtcacaaatggttaatgaaca | 1045 |
| Dp | 750  | GGGATGCTGCATCTGCTCGGGCCGGCCTTAATGACACAGTCATATGTTATGTAACA          | 809  |
| Qy | 1046 | ggtctctgccaggtcttgatggatgccaagaagaatcgaatctatggcatgtgatggaa       | 1105 |
| Dp | 810  | GCTCTCTCCAGAGCTGATGTGGATGCCACAGAAATCAGTATATGGCATATGATTTGAA        | 869  |
| Qy | 1106 | tggagacacccctcgaagatgacacaaatcttcctcgaagatcaacactctctgtgagaa      | 1165 |
| Dp | 870  | TGGGACCACTCTGAAAGTCAGCTCAATATCTCGAAGTCAACATTTCTTGAGAGA            | 929  |
| Qy | 1166 | accatcgcgaagcgctccttggaaatctcgcgcaataaacttcccttactgtgtcaaacactc   | 1225 |
| Dp | 930  | ACCATCGCCAGGCGCTTGGAATTCGCGCAATTACTTCTTACTGCTCAACACACT            | 989  |
| Qy | 1226 | tgaatgaccttgacaagttcttactgttttgtaatatctctcccaacaaatgatgga         | 1285 |
| Dp | 990  | TGATGACCTTGAGACAGTTTCTACTGTTTGTGATATCTTCTCCACCAACTGATGAGCA        | 1049 |
| Qy | 1286 | tggaaagctatgtaaaagttagaagcgtgtccagaggaaccccaacatagaatgaaata       | 1345 |
| Dp | 1050 | TGGAACTTATGTGAAGTAGACAGCTGCTCAGAGAACCCCACTACCAATGAAAAATA          | 1109 |
| Qy | 1346 | atgaaagcggaaagactatgatgatcttacttacttgaaatgtaatgtgtgcagt           | 1405 |
| Dp | 1110 | ATGAAGAAGCGGAAGACTATGATGATCTTACTGATTTCTGAATGATGATGTCAGCT          | 1169 |
| Qy | 1406 | ttagatgatgacaactctcctctcttataccaatctgcctcgaatgtgcgaagaactcta      | 1465 |
| Dp | 1170 | TTGATGATGACAACTCTCTCTCTTATTCACAAATTCGCTCAGTTGGCAAGAGCATCTTA       | 1229 |
| Qy | 1466 | aaacttggtaaccttaactatgtcgtctbaaagaagaagacggagactatgtccctaagcc     | 1525 |
| Dp | 1230 | AAACTTGGGTACATTTACATTTGCTGCTGAAGAAGAGACTGGGACTATGCTCTTATGCTC      | 1289 |
| Qy | 1526 | tcgcccccgatgacagaagttataaaagtaactatcttgacaatgtgcccacagcgatg       | 1585 |
| Dp | 1290 | TGCGCCCGGATGACAGAAGTTATAAAGTATATTGAAACAATGSCCTCAGCGGATTG          | 1349 |
| Qy | 1586 | gtaggaagtacaaaagaagctccgattatggtcalacagatgaaccttlaagactg          | 1645 |
| Dp | 1350 | GTAGGAAGTACAAAAAGTCCGATTTGTGSCATACACAGTGAACCTTTAGACTCTGTG         | 1409 |
| Qy | 1646 | aagcatctcagaatcgaatcaggaactcttggaccttacttactatggygaagttgagaca     | 1705 |
| Dp | 1410 | AAGCTATTCACCATGATTCAGGAATCTTTGGACCTTTACTTATATGGGAGATTTGAAGCA      | 1469 |
| Qy | 1706 | caacttgaattataattaaagataaagcaagagacatataacactcctccctcagaa         | 1766 |
| Dp | 1470 | CAGCTGTATTTATTTTAAAGATCAAGCAAGACGACCATTAACATCTACCTCAGGAA          | 1529 |
| Qy | 1766 | tcaactgatctgcctcttgaatcaagagatattacaaagttgtaaacatttgaag           | 1825 |
| Dp | 1530 | TCACGTAGTGTCCGTCCTTTGATTTCAAGAGATATTACAAAAGGTGTAAACATTTGAAG       | 1589 |
| Qy | 1826 | atttccaattctgcgcaggaagaatatacaataataatgtagacatgactgtagaagtg       | 1885 |
| Dp | 1590 | ATTTCCTCAATTCCTCCAGGAGAAATATTTCAAAATATTAATGAGACGTGTAAGAAGTG       | 1649 |

|    |      |   |      |
|----|------|---|------|
| OY | 1886 | gqccaactaatcagatccctcggtgcctgaacccgcgtatactctagtlttgtaatatbg      | 1945 |
| Db | 1650 | ggccmacctmaattcagatccctcggtgcctgaacccgcgtatactctagtlttgtaatatbg   | 1709 |
| OY | 1946 | agaagaagtctagcttcaaggagcttaattggcccctctccatctcgtcacaaagaatctyag   | 2005 |
| Db | 1710 | AGAAGATCTTAGGTTCAAGACTAATTGGCCCTTCCTCATTGTGCATCAAAGAATCTGTAG      | 1769 |
| OY | 2006 | atcaaagaaggaaaaccaataatgtcaagaagaagaaatgaacctgtttctcgttatgt       | 2065 |
| Db | 1770 | ATCAAAGAAGAAACCAATATATGTCAAGACAAGAGAAATGTCAATCCTGTITTTCTGTATTGG   | 1829 |
| OY | 2066 | atgsaaaccgaagcttgytlactcacagagaataatacaagctttctccccaatccagctg     | 2125 |
| Db | 1830 | ATGGAAGAACCGAAGCTGTACTTCCAGAGAAATATACAAAGCTTTCTGCCCAATCCAGCTG     | 1889 |
| OY | 2126 | gaagtcagcgtcttgagatccaggaattccaaagcctccaactcatgtacaagcatcaatggct  | 2185 |
| Db | 1890 | GAGTGCAAGCTTGAGAGATCCAGAGTTTCCAAGCCCTCCAAATCWTGACAGCATCAATGTGCT   | 1949 |
| OY | 2186 | atgtttcttgatagtttgcaattctgaatttgtttgttgatgaagtggaataacgylaacctc   | 2245 |
| Db | 1950 | ATGTTTTGTAGTGGTTTCCAGTTTGTCAAGTTTGTGTTTGTGATAGAGTGGCATACTGTGNATTC | 2009 |
| OY | 2246 | laagcatttgagacaagactgactcttcctctgtccttctctctgatalatacctccaac      | 2305 |
| Db | 2010 | TAAACATTTGGAGCACAGACTGACTTCTTCTTGCTGTTCTTGATATATACCTTCAAC         | 2069 |
| OY | 2306 | acaaatatgtgtctatgtgaagaacacatcaacctatcccatctcagagagaacgtgtctca    | 2365 |
| Db | 2070 | ACAAAATGGTCTATGAAGACACACTACCCTATTCCTCCATTCTCAGAGAAAACTGTCTTCA     | 2129 |
| OY | 2366 | tgtctgatgaaaaacccagagtgctatgtgattctcbgggtgtccaaacactcagaacttggaaa | 2425 |
| Db | 2130 | TGTGATGAGAAAACCAAGTGCTATGTGATTTCTGGGGTGCACAACTCAGACTTTTGGAAACA    | 2189 |
| OY | 2426 | gaagcatgacgcgcttacttgaagtttctcagtgtgtgaacaagaacactlgtatatltaag    | 2485 |
| Db | 2190 | GAGCATGACCGGCTTACGTGAAGTTTCTAGTTTGTATTCAGAGGGGGAGAGGAGCAGCG       | 2249 |
| OY | 2486 | agagcagttatgaatatatta-----gcatactgtctgagtaaaaaa                   | 2530 |
| Db | 2250 | ACTATTCGACCTGAGAGATATTCATGTAAGAGAGACAGCATACATCCATCTCTGCACA        | 2309 |
| OY | 2531 | atgscattgaaccaagaagatctcccaaatcaatgaagacacacctgacbtgsgnaaaagc     | 2590 |
| Db | 2310 | GTCGTGATTGAACCAAGAGAGCTTCTCCGAATAATTCAGAGACCCCTTAGCACTAGGCAAAAAGC | 2369 |
| OY | 2591 | aattaatgcccc-----   | 2603 |
| Db | 2370 | AATTTATATGCCACCAACAAATTCAGAAAATAGACATAGAGAAGACTGACCCCTTGTTTGAC    | 2429 |
| OY | 2604 | -----   | 2603 |
| Db | 2430 | ACAGAACCTATGCTMAAATACAAATATGTCCTCTAGTATTTGTTGATGCTTGGC            | 2489 |
| OY | 2604 | -----   | 2603 |
| Db | 2490 | GACAGAGTCTCTACTCCACATGCGGACTATTCCTTATCTGATCTCCAGAMGCCAATAATAGAGA  | 2549 |
| OY | 2604 | -----   | 2603 |
| Db | 2550 | CTTTTTCGATGATCCATCCACTGGAGCAATAGACAGTAATATACGCTGTCTGNAATGA        | 2609 |
| OY | 2604 | -----   | 2603 |
| Db | 2610 | CACACTTCAAGCCACAAGCTTCATCACAGTGGGACATGGTATTTACCCCTGATCAGGCC       | 2669 |
| OY | 2604 | -----   | 2603 |
| Db | 2670 | TCCAAATTAAGATTAAATGAGAAACTGGGAGACAAGTCAAGATCCTCTTGCTTGGGATAACC    | 2729 |

[illegible]

[illegible]

|                            |      |  |                            |
|----------------------------|------|--|----------------------------|
| Dd                         | 4890 | TCACACCTTGTTGGTAACCTCTGTAGAACCCACCGTCTACTGATCGGCTACCTTGCAATTTCACC  | 4949                       |
| Oy                         | 4616 | cccaagattggatgcacacgattgcccttcgaaggatgagaagtctcgtcgctgcagacag      | 4675                       |
| Dd                         | 4950 | cccagagttgggtgcaccacagattggcccttgagatggagsttctgggctgcagcacag       | 5009                       |
| Oy                         | 4676 | acctactgcagtcgagcgatlc   | 4699                       |
| Dd                         | 5010 | acctctactgaggggtggccacgtgc   | 5033                       |
| RESULT                     | 9    |  |                            |
| AR029098                   |      | AR029098   | 4334 bp                    |
| LOCUS                      |      | Sequence   | 38 From patent US 5859204. |
| DEFINITION                 |      | Accession  | AR029098                   |
| VERSION                    |      | AR029098.1   | GI:5941071                 |
| KEYWORDS                   |      |  |                            |
| SOURCE                     |      | Unknown.   |                            |
| ORGANISM                   |      | Unknown.   |                            |
| REFERENCE                  |      | Unclassified.  |                            |
| AUTHORS                    |      | 1 (bases 1 to 4334)  |                            |
| TITLE                      |      | Lollar,J.S.  |                            |
| JOURNAL                    |      | Modified factor VIII   |                            |
| FEATURES                   |      | Patent: US 5859204-A 38 12-JAN-1999;                               |                            |
|                            |      | Location/Qualifiers  |                            |
|                            |      | 1..4334  |                            |
|                            |      | /organism="unknown"  |                            |
| BASE COUNT                 |      | 1109 a   | 1126 c 1085 g 1014 t       |
| ORIGIN                     |      |  |                            |
| Query Match                |      | 66.9%; Score 3189.8; DB 9; Length 4334;                            |                            |
| Best Local Similarity      |      | 83.5%; Pred. No. 0;  |                            |
| Matches 3692; Conservative |      | 0; Mismatches 637; Indels 90; Gaps                                 |                            |
| Oy                         | 271  | atgcacaatagagctctccaactgcctctcttctctgcctcttgagatctcgtcttagtgcc     | 330                        |
| Dd                         | 3    | ATGCAGCTGATGACTCTCCACCTGTGCTTCTGTGCTCTTCCACTCGCGCTTAGTGCC          | 62                         |
| Oy                         | 331  | accagaagaatactactcctgggtgcattgaaactgtaatggactatatgcagaatga----t    | 387                        |
| Dd                         | 63   | ATCAGAGATPACTTACTCTGGCGCACATGTAACGTCTCTGGACTACCGCAAAGTAACTC        | 122                        |
| Oy                         | 388  | cctcgatgaactgcctcgtgcagcaaatctcctcctagatgccaataatcttccatc          | 447                        |
| Dd                         | 123  | CTCCGTGAAGCTGCAGCTGGACACACAGATTCTCTGTACAGGCCAAGGAGCTTCCGTTG        | 182                        |
| Oy                         | 448  | aacacctcaactcgtglaacaaaagaactcgttttgtagaattaocgyttocacttccaac      | 507                        |
| Dd                         | 183  | GSCCCGTCAGTCTCTGTACAAAAGAAGCTGTTGCTGATGATTCACGATCAACTTTTCAGC       | 242                        |
| Oy                         | 508  | atcgcctaagccaagggccaacccctcgatctggtctcgtcagtlccataccatccagctgaagtc | 567                        |
| Dd                         | 243  | GTTTGCAGGCCCGAGGCCACCATGATGGGTGCTGCTGATCCATCCAGGCTTAGGTT           | 302                        |
| Oy                         | 568  | tatgatagatgltgaatcaacttaaagaatagatgccatccctcgtgaactcttcaatgc       | 627                        |
| Dd                         | 303  | TACGACACGGTISGTCGTTTACCTGTAAGAACAATGGCTTCTCATCCGTTTAGTTCACCT       | 362                        |
| Oy                         | 628  | gttgtgtatccactactgtaaaagctctctgaaggagctgaatalgalatcatcagaacagtaaa  | 687                        |
| Dd                         | 363  | GTCGGCGTCTCTTCTGTGAATACTTCCGAAGGCGCTGATATMGAGATCCACACCGCAA         | 422                        |
| Oy                         | 688  | agggaagaagaatgatataaagctctccctcgtgtgaagccaatacatatgctcgaagtc       | 747                        |
| Dd                         | 423  | AGGGGAGAGGAAGATGAAGTAAAGTCTTCCCGGTAAAAACCAACTCATGTCGGCAGGTC        | 482                        |
| Oy                         | 748  | ctgaagaagaatggtccaatgacctgcctcgaacacacgtgctactactactatcttct        | 807                        |
| Dd                         | 483  | CTGAAGAANAATGTCACAAACGCTCTAACCACCATGTCTCATCTACTCTCATCTGCTCT        | 542                        |
| Oy                         | 808  | catgtgagacctgqlbaaagaacttgaattcaagqcccatcttgagacctctactglatgtga    | 867                        |



|    |      |   |      |
|----|------|---|------|
| QY | 3028 | cgctccattccctctcatctctagccctatttctcttagagaagatcagaagcaagagca        | 308  |
| Db | 2676 | GGTCCCTAATTCCTTCTACTCGAGCCCTTAATTTCTTAATCCGATGATAGAGGAAGGGCA        | 2735 |
| QY | 3088 | gaacctagaanaaacctctgtcagaagccctaatggaaccacaacttaactcttggaagtgcga    | 3147 |
| Db | 2736 | GAACCTGACACCAACTTCGTCTCAGCCAAATGAACCAAGAACTTACTTTTGGAAAGTCAG        | 2799 |
| QY | 3148 | catcatactgaccccaactaaagaatgatttgaactgcnaaagcctgagctatttctctgat      | 3207 |
| Db | 2796 | CATCACTATGGACCCACAGAAAGAGATTTGATCTGCAGAAAGCCGGACCTAATTTTCTGAT       | 2855 |
| QY | 3208 | gttgacctggaanaaagtctgacactcaagcgcctgaatttggaaccctctctgtctgccaact    | 3267 |
| Db | 2856 | GTTTACCTCGGAAAAAGATGTGCATCGCTGAGCTTATGGCCCTCTTGATGTCCGGCCGC         | 2915 |
| QY | 3268 | aacacactgaaacctgtgtcatctggagaagaatgaagaacagaatctgtcctgtcttctc       | 3327 |
| Db | 2916 | AACACCTGAACGCTGCTCAACGGTAGACAAAGTACCGTGCAGAAATTTCTCTGTTTTC          | 2975 |
| QY | 3328 | accaactcttgatgagaccaaaagcctgttacttcaactgtaanaataatggaanaaacctgcaag  | 3387 |
| Db | 2976 | ACTATTTTGTGATGAGACAAGAGACCTGCTACTTCACTGAAAAATGTGGAAAGCAACTGCCG      | 3035 |
| QY | 3388 | gtctcccgcaatatccgaagtgggaagatcccaactcttaagaagaatratatgtctctcatga    | 3447 |
| Db | 3036 | GCCCCCTGCCACTCGCAGATGAGAGAGACCCCACTCTAAGAAAGAACTATGCTTCCATGCA       | 3095 |
| QY | 3448 | atcaaatgtctacataatggaatacactaactctgtccttagtaatgtgtcctaagaatcnaaagat | 3507 |
| Db | 3096 | ATCAATGGCTATGTGATGAGATACACTCCCTGGCTTATGTAATGGCTCAGAAATCAAGGATC      | 3155 |
| QY | 3508 | cgatagatcatctgtctcagcatagggcgaagcaatgaaacaatccattctatcttaactgaatga  | 3567 |
| Db | 3156 | CGATGATATCGTCAACACATGGGACAGCAATGAAGAAATATCCATTCGATTAATTTTGGCGGA     | 3215 |
| QY | 3568 | catgtctcaactgtgacgaaaanaaaggagatataaagtgtgaactctacactctcatca        | 3627 |
| Db | 3216 | CACGTCTTCACTGATGACGGAAGAAAGAGAGAGATATAAATGCGGTGTACAAATCTTATACG      | 3275 |
| QY | 3628 | gggtcttcttgagacagatggaataatgttaccatccaaagctgtgaatttbgcgggtgtgaatgc  | 3687 |
| Db | 3276 | GGTGTCTTTTGAAGACAGTGAAGAAATGCTAACGTCMAAGTTGGAATTTGGCGAATTAAGAAAGC   | 3335 |
| QY | 3688 | cttatgtggagagcatctacactgtctggatggagaacactttctctgtgtgacagcaataag     | 3747 |
| Db | 3336 | CTGATTTGGCGAGCACTGCAAGGTGGAGTAGACACACTTTCCTGTGTACAGCAAGAG           | 3395 |
| QY | 3748 | tgctagaagtcacctgggaatctgtctctctgacacatataagatttcaagatttcaactcttca   | 3807 |
| Db | 3396 | TGTGAGGCTCCACGTGGGAATGCGCTCTCGACGCAATTAGAGATTTTTCAGATCAACACTTCCA    | 3455 |
| QY | 3808 | ggacaataatggaagatgtggccccaagactgtgccaagcttcaattatctcgatcaatcaat     | 3867 |
| Db | 3456 | GGACAGATATGAGACAGTGGGCCCAAAAGCTGGCCAGACTTCATTAATTCGGGATCAATTCAT     | 3515 |
| QY | 3868 | ggcctggagagccaaaggagacccttctcttgatcaagaatgtgactctcttggcaccaatgtat   | 3927 |
| Db | 3516 | GCCCTGGACACCAAGGATCCCACTCTCTGATCTAAGGTGAATCTGTGGCACCAATGTATC        | 3575 |
| QY | 3928 | attcaaggaatcagaagaccgaagggtggccgctcagaagatttctcagcctctacactctcga    | 3987 |
| Db | 3576 | ATTTCAGCGCATATGACCCAGGGGTGCCCGCAAAATTTTCCAGCGCTTACATCTCCAG          | 3635 |
| QY | 3988 | ttctatcatatgatatgattcttgatgggaagaagtgtgcagacttalcgaggaatctcaact     | 4047 |
| Db | 3636 | TTTATTCATCATATGATCACTCTTTGACGGGAGGAATGTGCACAGATTACCGAGGGATTTCCACG   | 3695 |
| QY | 4048 | ggaacctaatatgtctctctcttgacatgtgacttactctgggataaaacaataatcttt        | 4107 |
| Db | 3696 | GGCACCTTAATGGTCTTCTTTGGCAATGTGTGACGCATCTGGGATTTAAACACATATTTTTT      | 3755 |

| QY   | 4108                                 | aaacccaatattatgctccgataaaccggtttgaccccaactcaattatagatctgcagc     | 4167        |
|--|--------------------------------------|--|-------------|
| Db   | 3736                                 | AACCTCCGATGTGCGCTCGGACTACCGGTTTGCAACCAACATTAAGCATCCGACGC         | 3815        |
| QY   | 4168                                 | actctgcagatgattatgagctgctgattaaatagltgcagatgcatcttggaatg         | 4227        |
| Db   | 3816                                 | ACTCTGCCATGSGATTTATGGCGCTGTGATTTAAACAATTGCAGATCCCTCGGAATG        | 3875        |
| QY   | 4228                                 | gagagtaagaacataatcagatgacagataactgcttcacacctacttaccataatgltt     | 4287        |
| Db   | 3876                                 | CAGATTAAGGATATCGAGACTACAGATACAGGCGCTCTCCCACTTAAGCAATATATTT       | 3935        |
| QY   | 4288                                 | gcacccgtgtctcccttcaaaagcttgcacttcaactcccaaggagagatgactgctgaga    | 4347        |
| Db   | 3936                                 | GCCACCTGTGTCCTCTTCAAGACCCGACTTCACCTCCAGGGGCGGACGAATGCTTGCGGA     | 3995        |
| QY   | 4348                                 | cctcaggtgaaataatccaaaagatgctgtcaagtgaagtcacgaagaacataaagctc      | 4407        |
| Db   | 3996                                 | CCCCGGGTGACGACGCGCAGAGAGATGGCTGCAAGGTGACCTGCAGAMAGACGGTGAAGTTC   | 4055        |
| QY   | 4408                                 | acaaggaaactactcaaggagataaactctgtcttcaacgaatgatagtgaaagagctc      | 4467        |
| Db   | 4056                                 | ACAGGCATCACACACCCAGCGGCGTGAAGTCCCTGCTCACACAGATGTATGTGAAGAGTTC    | 4115        |
| QY   | 4468                                 | ctcaatcccaagatgcaagatgagccatgaatgagatctcctttttcgaataggcaagta     | 4527        |
| Db   | 4116                                 | CTCGGTCCAGTATGACAGGACGGCCGCGTGGAGCCCTGTTCTTCAAGAGCGCACACG        | 4175        |
| QY   | 4528                                 | aaggtttccagggaaatccaagactccttcaacactgtgtgtaactctctagaccacag      | 4587        |
| Db   | 4176                                 | AAGGTTTTCAAGGGCAATCAGAGACTCTCTCCACCCCGTGTGMAAGCTCTGGACCCCG       | 4235        |
| QY   | 4588                                 | ttactactgctactacttgaattcaaccaccagattggtgtgcaaccagattgcccagag     | 4647        |
| Db   | 4236                                 | CTGTTCACGCGCTACCTGAGATCCACCCCAAGAGCTGGCGCAGACATGCGCCGTGAG        | 4295        |
| QY   | 4648                                 | atggaagttctgtgctgtcgagcaaaagaaactctactga                         | 4686        |
| Db   | 4296                                 | CTCGAGGTTCTTGAAGTGTGAGGCAACAGGATCTCTAATGA                        | 4334        |
| RESULT 10  |                                      |  |             |
| LOCUS  | ARI10041                             | 4451 bp  | DNA         |
| DEFINITION   | Sequence 42 from patent US 6114148.  | PAT  | 14-FEB-2001 |
| ACCESSION  | ARI10041                             |  |             |
| VERSION  | ARI10041.1                           | GI:12826317  |             |
| KEYWORDS   | Unknown.                             |  |             |
| SOURCE   | Unknown.                             |  |             |
| ORGANISM   | Unclassified.                        |  |             |
| REFERENCE  | 1 (bases 1 to 4451)                  |  |             |
| AUTHORS  | Seed, B. and Haas, J.                |  |             |
| TITLE  | High level expression of proteins    |  |             |
| JOURNAL  | Patent: US 6114148-A 42 05-SEP-2000; |  |             |
| FEATURES   | Location/Qualifiers                  |  |             |
| source   | 1..4451                              |  |             |
| BASE COUNT   | 972 a 1568 c 1234 g 677 t            |  |             |
| ORIGIN   |                                      |  |             |
| Query Match  |                                      |  |             |
| Best Local similarity 74.8%; Pred. No. 0;                        |                                      |  |             |
| Matches 3266; Conservative 0; Mismatches 1101; Indels 0; Gaps 0; |                                      |  |             |
| QY   | 328                                  | gccacccaagaatactactcctgtgtgcagtgaaactgtcatggaactatatagtcaaatgtag | 387         |
| Db   | 84                                   | GCCACCGCCGCTACTACTGCGGCGCGTGAGCGTGCCTGCGGACTCATCATCAAGAGCGAC     | 143         |
| QY   | 388                                  | ctgggtgagctgcctgtgagcgaagaattctctctctagagtgcacaaatctttccatc      | 447         |
| Db   | 144                                  | CTGGGCGAGCTCCCGTGTGAGCGCCGCTTCCCGCCGCGTGCACAGAGCTTCCCTTC         | 203         |



QY 448 aacacctgaagctgctgtacaaaagaactctgttttgtaattcagagtttccactttcaac 507  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 204 AACACCAGCGTGGTGCACAAAGAAACCCTGTTCTGGAGTTTACCGACACCTTTTCAAC 263  
QY 508 atcgttaagccaagaagccaccctgtgatgtgtctgttaagtcctaccacacagagctgaagt 567  
||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||||  
Db 264 ATTGCCAAGCGCGCCGCCCTTGATGGCGCTGCTGGGGCCCACTTCAGCGCGGAGGTG 323  
QY 568 tatgtacagtgatcattacacattagaacaatvgcttccactctgtcagttcattatgt 627  
||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||||  
Db 324 TACACACCGGTGTATCACCCTGAAGAACATGTGCGCAGCCACCCTGACGCTTCACGCC 383  
QY 628 gtgtgtatctactatggaagagctctgtgagagctgaataatgtatgtacagacagca 687  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 384 GTGGGCTGAGACTGTGAAGAGCCAGCAGAGGGCCGCAAGTACGACACAGCTTCCAG 443  
QY 688 aggtagaagaagaatgtataagctcttcctgtgtgaagccaatacatatgtctgagctc 747  
||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||||  
Db 444 CGCGAAGAGGAGGACGACAGGTGTCCGGGGGGAGCCACACCTTACGTGTGGCAGGTG 503  
QY 748 ctgaagaagaatgttccaatgtgctctgacccactgtgtcttactctaatatcttct 807  
||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||||  
Db 504 CTTAAGGAGAACGGCCCTATGCGCAGCAGACCCCTGTGCTGACCTACAGCTACTAGC 563  
QY 808 catgtgacctgtgtaaaagacttgtaattcagagctctatgtgagccctactagatgtaga 867  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 564 CACCTGAGACTGTGAAGAGTGTGAACAGCGGCTGTATCGCGCGCTGTGGTGTGCG 623  
QY 868 gaaagagctctgtgccaagaagaagaacagacactctgcacaatttatacacttttgtct 927  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 624 GAGGGCAGCCTGGCCAGAGGAAACCCAGACCTTGCAAAAGTTTCACTTCGCTTTGGCC 683  
QY 928 gtatttgatgaagaaagtltgacctcagaacaagaactccttgatgtcagagatagg 987  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 684 GTGTTCCACAGGAGGGAAGACTGTGACAGCAGGACTTAAGAACACCTGATGACGAGACGC 743  
QY 988 gatgtcatctgtctgtgagctgtgagctaaatgtcacacagttcaatgttatgttaacagg 1047  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 744 GACGCCCGCAGCGCCCGCTGGCCCAAGATGTGACACCCGTTAAAGGCTTACGTAAACCGC 803  
QY 1048 tctctgcaagctctgtatgtgatgtcagcaggaatcattctatgtcagtgtatgtgaatg 1107  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 804 AGCCTGCCCGGCTGATCGGCTGCGCACCGCAAGCGGTATCTGCACTCATGGGCATG 863  
QY 1108 ggcacacacctgtgaatgtcacaatttccctcgaaggtcacaacatttctgttgaaggac 1167  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 864 GGCACACACCCCTGAGTGCACAGCATCTTCCTGGAGGGCCACACCTTCCTGTCGCACAC 923  
QY 1168 catgccaagcgctccttggaatctcgcacaataacttctcttactgttcaaaactcttg 1227  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 924 CACCGCCAGGCGACCTGTGAGATACGCCCATACCTTCCTGACTGACCCAGACCTTCTG 983  
QY 1228 atggaacttgacaggttcttactgtttgttcatactctctcccaacaactgtacgtcatg 1287  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 984 ATGGACTTAGGCCAGTTCGTGCTGTCTGCACTTCAGCAGCCACACGACGAGGGCATG 1043  
QY 1288 gaagcttatgtcaaatgtagaagctgtctcagaggaacccaactacgaatgtgaaataat 1347  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1044 GAGGCTTACCTGAAGGTGAGCAGCTGCCCCGAGAGGCCCAAGCTGCGCATGAAGAACAC 1103  
QY 1348 gaagaagcgaagaactatgtatgtacttactgtatctcgaatgtgagtgtgtcagttc 1407  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1104 GAGAGGCGCCAGGACTGTGACGACGACCTTACCGACGACGAGATGTGTACGGCTTC 1163  
QY 1408 gatgatatacaactctctcttcttataccaatctgtcagttgttcgaagaagactctcaaa 1467  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1164 GAGCAGCAACAAAGCCCGAGTTTATCCAGATCCGCAAGCGGTGGCCAAAGAACCCCTAAG 1223  
QY 1468 acttgggtacattacattgtctgtgaagagagagactgtggaattgtcctctagttctc 1527  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1224 ACCTGGGTGACTACATTCGCCGCCGAGAGGAGAGACTGTACGCCCGCTAGTACTG 1283

QY 1528 gccccgatgacaaagtattataaagttaatttgaacaaatggccctcagcgatttgt 1587  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1284 GCCCCGACGACCCGACGCTTACAAAGCCAGTACTGAAACAAAGGCCCCCGACGCAATG 1343  
QY 1588 aggaatatacaaaaagtcctgaattatgtcatacacaatgaatacctttaagactctgtga 1647  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1244 CGCAAGTACAAAGAGGTGGCTTCATGGCTTACACCGACGAGACTTTTCAAGAACCCGGAG 1403  
QY 1648 gctatcagaatgaatacagaacttgggaccttacttataatgggaagtgtgagaaca 1707  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1404 GCAATCCAGCAGATCGCGGATCTCGGCCCTTCGTGTACGGCGAGTGTGGGAGCCAC 1463  
QY 1708 ctgtgttatattataagaatcaagcaagcagacacataaactcttaccctcagaagatc 1767  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1464 CTGCTGATCATTTCAAGAACCAAGCCAGCAGGCCCTTACAAACATCTTACCCACGGCATTC 1523  
QY 1768 actgatgtccgttccctgttattcaaggagatttacaagaagtgtgaaacatttgaagat 1827  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1524 ACCGACGTGGCGCCCGCTGTACAGCGCGCGCTGCCCCAAGGGCGTGAAGCCTGAAGGAC 1583  
QY 1828 ttccaattctgcagaagaagaatatcaaatataatgtgacagtgtcgtgaagaatgg 1887  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1584 TTCCCATCTCCGCGGGGAGATCTTCAAGTACAAAGTGGACCTGTGAGCGAGACGCGC 1643  
QY 1888 ccaactaatcagaatcctcsgtgcctgacccgctattactcttaatttctgttaatatagg 1947  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1644 CCCACAAAGACGACCCCGCGCTGACCCGCTACTACAGACACTTTCGTAACATGGAG 1703  
QY 1948 agaatctagcttcaaggaactatgtgcccctcctcctcactctgtcacaagaatctgtagat 2007  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1704 CGGACCTGGCGCTCGGAGCTGATGGGCCCTGCTGATCTCTTCAAGAAAGAGGTGAGC 1763  
QY 2008 caaagagaacaacagataatgtcagaagaagaatgtcactgtttctgtatltgtat 2067  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1764 CAGCGCGCAACCAAGATCATGAGCGACAGCGCAACGATGATCTGTTCAGCTTTCGAC 1823  
QY 2068 gagaacgaagctgtatcctcacaagaataataaagcttctcccaatccagctgtgga 2127  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1824 GAGAACCGACAGCTGTGTCTGACCGAGAACATCCAGGCGTTCCTGCCCCAGCCGCTGGC 1883  
QY 2128 gtgcagcttgaggaatcagaatgttcaagcctcacaacatcaltgacaagcatatgtctat 2187  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1884 GTGCACTGGAAGATCCCGAATTCAGAGCCGACGAACATCATGACAGATCAAGGGCTAC 1943  
QY 2188 gttttgatagtttgcagttgtgtcagttgtgttcgaatgaagtgtgcatactgttactcta 2247  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1944 GTTTCGACACGCTGTGAGCTGAGCTGTGCTGCTGATGAGTGGCTTACTGTATCTCTG 2003  
QY 2248 agcatlggagacagactgtacttcttctgtcttcttctgtatatacttcaaacac 2307  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 2004 AGCATGCGCGCCGACCGAGCTTCTGAGCGTGTCTTCGCGGGTATACCTTCAAGCAC 2063  
QY 2308 aaaaigtctatgaagacacactcaccatcttccactctcagaagaagaactgttctatg 2367  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 2064 AAGATGTGTATGAGAGAACCCCTGACCTGTTCCTTCCTCGGCGAGACTGTTCATG 2123  
QY 2368 tcgaatgaaaaacccaggtctcatgtatctcgggtgtgccaaactcagaacttctgagaaga 2427  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 2124 TCTATGAGAACCCCGCGCTGTGGATTCTGGGCTGCCAACACAGCATTCCTCGCAACGCG 2183  
QY 2428 ggcattgacgcttactgaaggttcttctgtgtgtacaagaacactgtgtattatcagag 2487  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 2184 GGCATGACTGCGCTGTGAAGTCTTCCAGCTGCGCAAGAACACCGCGACTTACTACGAG 2243  
QY 2488 gacagtatgaagaatatctcagaactctgtgtgttaaaaacatgtcattgtgaacaga 2547  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 2244 GACAGCTTACGAGGACATCTCGGCTACTCTGTGTGTCAAAGAACAGCATGAGACCCCGC 2303  
QY 2548 agcttctccagaatctcaagaacacctagcactagcgaagaatlttaatgtcaccaca 2607  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 2304 TCTTTCTCCAAAACTCTCGCGCACCCGACGAGCTGAGAAAGAGTTCACAGGCCACCCCC 2363  
QY 2608 ccagttctgaaacgcatacgaaggaataactcgttactactcttcaagtcagatcaagag 2667



|    |      |                                   |                                |      |
|----|------|-----------------------------------|--------------------------------|------|
| Db | 2364 | CCCCGCTGGAAGCGCGCACACGCGAGATCA    | CCCCACCCCTCCAAAGCAGCAGGAG      | 2423 |
| Qy | 2668 | gaattgaactatgatgacatcatcagttgaatc | gaagaagatttgaactt              | 2727 |
| Db | 2424 | GAGATCGATGACGACGACCATCAGCTTGAG    | ATGAAGAGAGACTTCACATCTTAC       | 2483 |
| Qy | 2728 | gatgagatgaaaaatcagagcccccagcttca  | aaagaanaacagacatctt            | 2787 |
| Db | 2484 | GAGACGAGCAGAAACCAAGCCCCGCTCTTC    | CAAAAAGAAAACCGGCACACTTCAATC    | 2543 |
| Qy | 2788 | gctgcagtgaagagctcttgatataatgagta  | gtagtcccaatgcttcaagaac         | 2847 |
| Db | 2544 | GCCCGCCGGGAGCGCTGTGGGACTACGAC     | AGACAGACGCCCCACACTCTCTGGCAAC   | 2603 |
| Qy | 2848 | agggtcgaagtggcagtgctccatcagtaaga  | agtgttctcaggaatttactgat        | 2907 |
| Db | 2604 | CGCCCGCAGAGCGGAGCGTCCCAAGTTCA     | AGAAAGTGCTTCCAGAGATTCAACGAC    | 2663 |
| Qy | 2908 | ggtccttctactcagcccttataccgttgaga  | actaaatgaacatltggactcctggg     | 2967 |
| Db | 2664 | GGCAGCTTACCCAGGCCCTGTACCGCGGAG    | CTGTAAGCAGCACCCTGGGCTGCTGGC    | 2723 |
| Qy | 2968 | ccatatataagaagcagaagtgaagataata   | tcatgttaacttcaagaatcagcctct    | 3027 |
| Db | 2724 | CCCTACATCCCGCGGAGAGTGGAGCAACAT    | CAATGATGACCTTCCGCAACCAACCTCC   | 2783 |
| Qy | 3028 | cgctccatctctcttacttactcagcttact   | cttcttgaaggaagcagaagcagaagca   | 3087 |
| Db | 2784 | CGGCTCTACTCTTCTACTCTCTCTGATCA     | GCTACGAGAGAGGACGAGCGGAGCGCG    | 2843 |
| Qy | 3088 | gaacctagaanaaaacttctgtaagacctta   | atgaacaaacttacttcttgaagaatgca  | 3147 |
| Db | 2844 | GAGCCCGCAGAAACTTGTAAGCCCAACGAG    | ACTAAGACCTTCTTGAGAGGTGGAG      | 2903 |
| Qy | 3148 | cataatataggcaccactaaagatgagttga   | ctgcaagagcctgggtctattctctgat   | 3207 |
| Db | 2904 | CACACATGCGCCCGCCACAGAGAGAGTTG     | CGACTGCAAGGCTGGGCTCTACTTTCAG   | 2963 |
| Qy | 3208 | gttgacctgaanaaagtgtgtaacctcagct   | gatgagccctctgtctgcacacact      | 3267 |
| Db | 2964 | GTGGAGCTGGGAGAGAGGTGCACAGCGGCT    | GTGATCGGCCCTCTGTGTGCCACAGC     | 3023 |
| Qy | 3268 | aacacactgaaccctgtctcatctgagagaca  | agtgtgacatgacagaaatltgtctgtttc | 3327 |
| Db | 3024 | AACACCGCTGAACCCCCCGAGGGAGGCA      | AGTGTGACTGTGACAGAAATTGCCCTGT   | 3083 |
| Qy | 3328 | accactttgatgagaccaaagtgtgtacttca  | ctgtaaaataatgaagaagaactgcag    | 3387 |
| Db | 3084 | ACCATCTTCGACGAGACTTAAGGTGATCTT    | CACCGAGAACATGAGCGCAACTGCGC     | 3143 |
| Qy | 3388 | gctccctgcaaatatccagaatggaagatcc   | caactttaaagaaatlatctgtccatgca  | 3447 |
| Db | 3144 | GCCCCCTGCAACATTCAGATGGAAGATCC     | CACTTCAAGGAGAACTTACCGCTTCCAC   | 3203 |
| Qy | 3448 | atcaaatgtctacaaatgatacactactgtg   | ctagttaatggtccagatcaaaagat     | 3507 |
| Db | 3204 | ATCAAGCGGTACATCATGATGACCTGCGG     | CGCTGTGTGATGCGCCAGAGCAAGCGCA   | 3263 |
| Qy | 3508 | cgtatgtatctgtctcagcaatgggcagcaa   | tgaaaaaatccatctcatcttcaatgtga  | 3567 |
| Db | 3264 | CGGTGATACCTGTCTGTATGTGGGAGCA      | AGCAAGAACATCAACAGCATCTTTCAG    | 3323 |
| Qy | 3568 | catgtatctactgtcaagcaaaaaaagagta   | ataaatggtgcacttacaatctcatcca   | 3627 |
| Db | 3324 | CACCTTTTACCGTGGCAAGAGAGAGATG      | CCCTGTACAACTTGTACCC            | 3383 |
| Qy | 3628 | gtgttttttagacagatggaatgttaccatca  | aagcttgaaatctggcgggttggaatgc   | 3687 |
| Db | 3384 | GGGTGTTCAGAGCTGTGATGCTGCCAG       | CAAGCGCGGAGATGTGGCGCTGGAGTGC   | 3443 |
| Qy | 3688 | cttatgtgcagatctacatgtctgagtcaca   | cttcttctgtgtacaagaataag        | 3747 |

---

|    |      |                                 |                                |      |
|----|------|---------------------------------|--------------------------------|------|
| Db | 3444 | CTGATGGCGAGCACTGTGCAGCGCGCAT    | AGAGCAACCTTCTCTGTGTACAGCAACAG  | 3503 |
| Qy | 3748 | tgtaagactcccttggaatggtcttgcacat | atagaatttcaagatccacttca        | 3807 |
| Db | 3504 | TGCCAGACCCCTTGCGCATGGCCAGCGG    | CCACATCCGCACTTCCAGATCACCGCCAG  | 3563 |
| Qy | 3808 | ggacatatgacagtggtgccccaaagtctgc | agacttcaatcttaltccgataatcaat   | 3867 |
| Db | 3564 | GGCCAGTACGGGCATGGGCTCCCAAGCT    | GGGCCCTCTGCACTACAGCGGAGCATCAAC | 3623 |
| Qy | 3868 | gcttgagcacaagaagccttcttcttgatca | agtgatgtctgttgcccaatgat        | 3927 |
| Db | 3624 | GCTGTGTGACCAAGAGCGCTTCTCTGTG    | ATCAAGTGGACCTGTCTCCCATATTC     | 3683 |
| Qy | 3928 | attcagcatcaagaccaggtgtccgtcaga  | agtcttccagcctctacatctcag       | 3987 |
| Db | 3684 | ATCCAGCGCATACACCCAGGCGCCCGC     | CAAGATTCAAGAGCTGTACATCAGCCAG   | 3743 |
| Qy | 3988 | ttatcatatgtatagttcttgaatggaga   | agtgtgcagacttalcgaagaatccact   | 4047 |
| Db | 3744 | TTCATCATCATTACTCTTACAGCGCAAG    | AAAGTGGACACTACCGCGCAACGACAC    | 3803 |
| Qy | 4048 | ggaaccttaatgtcttcttcttggaatgt   | gatattcatcttggaataaacaatatltt  | 4107 |
| Db | 3804 | GGCACCCTGATGTGTCTTCTGCAAGT      | GGACAGCAGCGCATCAACCAACATCTTC   | 3863 |
| Qy | 4108 | aaccttcaatatattgtctgcagatac     | tcggttgcaaccacatctatagcatcagc  | 4167 |
| Db | 3864 | AACCCCGCATATGTGCCGCTACATCCG     | CTGCAACCCCACTACAGATCCGCAAC     | 3923 |
| Qy | 4168 | actctgcatagtatgtatggtgtgtat     | taaatagttgcagaatgcatlyggaat    | 4227 |
| Db | 3924 | ACCTGTGCGATGAGACTATGGGCTGCA     | ACCTGAGACACTGACAGATCCCTTGCG    | 3983 |
| Qy | 4228 | gagagttaagaacataatcagatgacaga   | gttactgtcttacttacttaccataltt   | 4287 |
| Db | 3984 | GAGACCAAGGCGCATAGCAGCGCCCA      | GATCAACCGCTTCACTTCCAAACATGTC   | 4043 |
| Qy | 4288 | ggcaccctgtctccttcaaaagctgcact   | ccaactccaaggaggaatgctgtgaga    | 4347 |
| Db | 4044 | GCCACTGTGAGGCCCGCAAGCAAGGCG     | CGCTGTGACCTGCAAGGCGCAGCAAG     | 4103 |
| Qy | 4348 | cctcagtgtaataatccaaaagtgtgctga  | agtgtgaaatccagaagaatgaanaa     | 4407 |
| Db | 4104 | CCCCGAGTGAACCAACCAAGAGATGCT     | GCAGGCTTCCAGAAAACCATGAAGTG     | 4163 |
| Qy | 4408 | acagagtaactactcagaggagtaaat     | ctctgtcttaccagcaatgtatgtgaag   | 4467 |
| Db | 4164 | ACTGSCGTGACCAACCGAGGCGTCAAG     | AGCCTGTGACCAAGCATGTACGAAGGATTC | 4223 |
| Qy | 4468 | ctcatctccagcagtcagaatggtccatag  | tgcagtgactccttlttcaagaatgcaag  | 4527 |
| Db | 4224 | CTGATTCAGAGCAGGCAAGGAGCGCC      | CAACCATGAGACCTGTCTTCCAAAAG     | 4283 |
| Qy | 4528 | aaggttttccaggaatcaaatcaactcttca | caacctgtgtgaactctctagaccac     | 4587 |
| Db | 4284 | AAGGTGTTCCAGGCAACCAAGCAGCTT     | CAACCGGCTGTGAACAGCCTGGAAC      | 4343 |
| Qy | 4588 | ttactgactcgttactcttgcgaatltca   | cccccaaggttggtgtgcaacagatgt    | 4647 |
| Db | 4344 | CTGCTGACCCGCTTACTCTGCACTCC      | CCCAAGACTGTGCACCAAGATGCGCT     | 4403 |
| Qy | 4648 | atggaggttctggtctgcagagcacagga   | ccctctctactcagagcg             | 4694 |
| Db | 4404 | ATGGAAGTGCTGTGGCTGCGAGGCG       | CAAGACCTGTACTTAAGCGGCG         | 4450 |

RESULT 11  
AR003710 9009 bp DNA PAT 04-DEC-1998  
LOCUS  
DEFINITION Sequence 1 from patent US 5744446.  
ACCESSION AR003710  
VERSION AR003710.1 GI:3964969



|   |                                     |  |                 |
|---|-------------------------------------|--|-----------------|
| Qy  | 2130                                | tttgaatgctgcagctgtgcattgttcttgatgaagggcgatcactgtaattcctaag         | 2249            |
| Db  | 2070                                | tttttatattgttgacatttctcagtttgattgttgacatgagggccattgctgatacttcttaag | 2129            |
| Qy  | 2250                                | catlgaagacacagacttcttctctgtctctctctctctgatalacttcaaacaca           | 2309            |
| Db  | 2130                                | catttgagacacagactgacttcttctctgtctctctctctcttgatattacttcaaacaca     | 2189            |
| Qy  | 2310                                | aatgctctatgaagacacactcaaccttattccattctcgaagaaactgtcttcatgct        | 2369            |
| Db  | 2190                                | aatgctctatgaagacacactcaacctattccattctcgaagaaactgtcttcatgct         | 2249            |
| Qy  | 2370                                | gatggaaaaccagatctctatgattctctgggtgtgccaaactcaagacttccgaaacagag     | 2429            |
| Db  | 2250                                | gatggaaaaccagatctctatgattctctgggtgtgccaaactcaagacttccgaaacagag     | 2309            |
| Qy  | 2430                                | catgaacgccttactgaaggtcttctaagtcgtgtacaaagaacactggtgatatcagaaga     | 2489            |
| Db  | 2310                                | catgaacgccttactgaaggtcttctaagttgtgtacaaagaacactggtgattattacagaga   | 2369            |
| Qy  | 2490                                | cagttatgaagatattcagaactctgtctgtgataaaacaatgcatlgaaccaagaag         | 2549            |
| Db  | 2370                                | cagttatgaagatatttctcagacttctgtctgtgataaaacaatgcatlgaaccaagaag      | 2429            |
| Qy  | 2550                                | ctctcccaagaatcaagaacaccctagacactgagcgaagaacattaatgcaacocca         | 2607            |
| Db  | 2430                                | cttctcccaagaattcaagaacacccttagacactgagcgaagaacattaatgcaacocca      | 2487            |
| RESULT 12   |                                     |  |                 |
| AR029065  | LOCUS                               | AR029065   | 9009 bp DNA     |
| DEFINITION  | Sequence 1 from patent US 5859204.  |  | PAT 29-SEP-1999 |
| ACCESSION   | AR029065                            |  |                 |
| VERSION   | AR029065.1                          | GI:5941038   |                 |
| KEYWORDS  |                                     |  |                 |
| SOURCE  | Unknown.                            |  |                 |
| ORGANISM  | Unknown.                            |  |                 |
| REFERENCE   | Unclassified.                       |  |                 |
| AUTHORS   | 1 (bases 1 to 9009)                 |  |                 |
| TITLE   | Loi1ar,J.S.                         |  |                 |
| JOURNAL   | Modified Factor VIII                |  |                 |
| FEATURES  | Patent: US 5859204-A 1 12-JAN-1999; |  |                 |
| Source  | Location/Qualifiers                 |  |                 |
|   | 1..9009                             |  |                 |
|   | /organism="unknown"                 |  |                 |
| BASE COUNT  | 2853 a 1907 c 1844 g 2405 t         |  |                 |
| ORIGIN  |                                     |  |                 |
| Query Match 49.0%; Score 2336.4; DB 9; Length 9009;           |                                     |  |                 |
| Best Local Similarity 100.0%; Pred. No. 0;                    |                                     |  |                 |
| Matches 2337; Conservative 0; Mismatches 1; Indels 0; Gaps 0; |                                     |  |                 |
| Qy  | 270                                 | catgcaaatagagctctccacactgtctcttctgtgaccttttgagattcgtcttaagtc       | 329             |
| Db  | 150                                 | catgcaaatatagagctctccacactgtctcttctgtgaccttttgagattcgtcttaagtc     | 209             |
| Qy  | 330                                 | caccagaagaatactaccctcgggtgagcagatgaaatcgaatgagactatagcaagagact     | 389             |
| Db  | 210                                 | caccagaagaatactaccctcgggtgagcagatgaaatcgaatgagactatagcaagagact     | 269             |
| Qy  | 390                                 | cagtgagctgcctcgtgagacgcaagaatttccctcctagagtgccaaaactcttccatcaa     | 449             |
| Db  | 270                                 | cagtgagctgcctcgtgagacgcaagaatttccctcctagagtgccaaaactcttccatcaa     | 329             |
| Qy  | 450                                 | caacctcagtcgtgtacaaaagaactcgtgttctgtagaatcacggttccactttccaacat     | 509             |
| Db  | 330                                 | caacctcagtcgtgtgtacaaaagaactcgtgttctgtagaatcacggttccactttccaacat   | 389             |
| Qy  | 510                                 | cgcctaagcgaagggcaacctgtagatggtgtctgttaagtcctaccatccagggctggaagtta  | 569             |

|    |      |  |      |
|----|------|--|------|
| Db | 390  | CGCTAAGCCAAAGGCCACCTCGTGATGGGTCGTGATGGTCTACCATCAGCTGAGGAGTTTA    | 443  |
| Qy | 570  | tgaacagtggtcaattacacttaagaacatggtctccactcctgttcagttcattatgctgtc  | 629  |
| Db | 450  | TGATACAGTGTGTCTATTCACCTTAAGAACATGAGCTTCCATCCTGTGACGTTCAATGCTT    | 509  |
| Qy | 630  | tgtgtatctactggaaacctctctgaaggagctgtaatttgltgtcagacagcgaag        | 689  |
| Db | 510  | TGGTGATCTCTACTGGAAACCTTCTGAGGAGCTGTAATGTGATGATCACACAGTCAAG       | 569  |
| Qy | 690  | ggaagaagaatgatataaagtcctccctggtgtaaaagccatacatgltctgcagtcctc     | 749  |
| Db | 570  | GGAAGAAAGAAATGATATAAATCTTCCCTGGTGGAAACCATACATATGTCTGGCAGAGTCT    | 629  |
| Qy | 750  | gaagagaatggtccaaatggtcctctgacccaactgtgctctactcaactcaatccttcca    | 809  |
| Db | 630  | GAAAGAAATGATGATGATGCTGACCTGTGACCCACAGTGTCTCTACTCATATCTTTTCA      | 689  |
| Qy | 810  | tgttgacctgtgtaaaagacttgtaattcaggccctcaattgagccctactagatgtaaga    | 869  |
| Db | 690  | TGTGACCTCTGTAAATAACTTGAATTCAGGCTCATTTGAGGCCCTTAATGATGTAAGAGA     | 749  |
| Qy | 870  | aggaagctgtgccaagaaagacacagacacctgtgacaaattataactacttctgtcgt      | 929  |
| Db | 750  | AGGGAGTCTGGCCAGGAAAGACACAGACACTTGCACAAATTTATACTAATTTTTCGCT       | 809  |
| Qy | 930  | atttgatgaaggaaagatttgacctcgaagaacaaagaacccctctgatgacgaatagga     | 989  |
| Db | 810  | ATTTGATGAAGAGGGAAGGTTTGGCAGCTCAGAAACAAAGAACTCTTGATGACAGATAGGA    | 869  |
| Qy | 990  | tgtctgacatctgtctcggtcggtcgtcctaataatgacacagacaaatglttbtgtaaaagtc | 1049 |
| Db | 870  | TGCTGCACTGTGCTCGGGCTGTGGCTTAAATAGCACAGATCAATGGTTATGTAACAGCTC     | 929  |
| Qy | 1050 | tctctcaggtctgatttgatggtccacaggaatcaactatgcatatgcatgattgtaagatgg  | 1109 |
| Db | 930  | TCTGCCAGTCTGATTGTGATGCCACAGAGAAATGACGTCTTGCGATGTGTAATGG          | 989  |
| Qy | 1110 | caccactcctcgaagtgcactcaatattcctcgaaggtccacacattctctgtgaagaacca   | 1169 |
| Db | 990  | CACCACCTCTGGAAGTGCACACTAATTTCTCGAAGGTCCACACTTTCCTTGAGAACCA       | 1049 |
| Qy | 1170 | tctgcgaagtgctcttgaaatcttcgcgaataaacttctctactgtctaaaacctctgat     | 1229 |
| Db | 1050 | TCGCGACGCTCTCTGTGAAATCTCGCCAAATATCTTCTACTGCTCAACACCTCTGAT        | 1109 |
| Qy | 1230 | ggaacctggaacatttctactgttctgcatatctctccacacaatgtagcatgga          | 1289 |
| Db | 1110 | GGACCTTGGACAGATTCTACTGTTTGTGCTATCTCTTCCACCAACATGATGCGATGGA       | 1169 |
| Qy | 1290 | agcttatgtcaaaagtacagacgtgtccagaagaaacccaactacgaatgaaanaataaga    | 1349 |
| Db | 1230 | AGAAAGCGGAAGACTATGATGATGATCTTCTGTAATTCGTAAGATGAGTGTGAGGTGTGA     | 1289 |
| Qy | 1410 | tgatgaacactctcccttcccttcccaaatctcgctcaagtgtgccaagaagactccctaaac  | 1469 |
| Db | 1290 | TGATGACAACTCTCTTCCTTTTCCAAATTTGCGTCAAGTGTGCCAAGAACATCCTTAAC      | 1349 |
| Qy | 1470 | ttgggtacattacatgtctgtctgaagaaggaagactgtagactatgctccttactctcgc    | 1529 |
| Db | 1350 | TTGGGTGATTAATGATGCTGCTGAAGAAGAGACGTGGACATGTGCTCCCTTAATGCTCGC     | 1409 |
| Qy | 1530 | cccccgatgaacagaagttataaagttcaaatattgaaacaatggtcctctcagggatgtgtag | 1589 |
| Db | 1410 | CCCCGATGACAGAAGTTATTAAGTCAATATTTGAAACAATGGCCCTCAGCGATTTGATG      | 1469 |
| Qy | 1590 | gaagtaaaaaagtcggatttttctgcatcacacagatgaacactttagactctgtgaac      | 1649 |
| Db | 1470 | GAAATCAAAAAAGTCGATTTATGTGCATACACAGATGAACCTTTAAGACTCTGTGAAC       | 1529 |

1650 taccagcatgaatcagaatcttggaccttacttattgaggaagtgtgagacacact 1709  
1530 TATTACAGATGAATCAGAAATCTTGGGACCTTTATGGGAAATTTGGAGACACACT 1589  
1710 gtgattatattaaagaatcagaacagacacataaacttaccctcagcgaatcac 1769  
1590 GTTGATTATATTAAAGATCAAGACAGACATATACATCTACCCCTCAGCGAATCAC 1649  
1770 tgatgtccgtcttctgattcacaagagattaccaaaagtgttaaacatttgaagattc 1829  
1650 TGATGTCCGCTCTTATTTCAGAGATTACCAAAAGGTGTAATAACATTTGAAGGATTT 1709  
1830 tccaatcttcgcaagagaataattccaataataaaggacagtgagcttgaagaatggcc 1889  
1710 TCCATTCTGCAGAGAAATATTCAATATTAATGAGAGGTGACTGTAGAGATGGGCC 1769  
1890 aactaaatcagaatcctcgtgctgacccgtatctacttacttacttctgtaataatgagag 1949  
1770 AACTAAATCAGATCCTCGGTGCGCTGACCCGCTATTACTCTAGTTGCTTATATGAGAG 1829  
1950 agatctagcttcaagactcatttggccctccctcaatctgctcaagaagaatctgtatca 2009  
1830 AGATCTAGCTTCAGGACTCATGTGCCCCCTCTCTCATCTGTACAAAGAACTGTAGATCA 1889  
2010 aagaggaaccagataatgtcagacacagagagaatgtacatctgttcttattatgtga 2069  
1890 AAGAGGAACCAATATATGTGACAAAGAGAAATGTATCTGTTCTGTATTTGATGA 1949  
2070 gaaccgaagcttctgacccctcagagaataatacaagccttctccccaatcagcttgaat 2129  
1950 GAACCGAAGCTGGTACTCTCAGAGAAATATACAACGCTTTCTCCCAATCAGCTGAGT 2009  
2130 gcaagcttgaagatccagaagtccaagcctccaacataatgcaagaatcaatgtgctatgt 2189  
2010 GCAGCTTGAAGATCCAGAGTTCCAAAGCCTCCAAATCATCATCACAGCATGCTATGT 2069  
2190 ttctgattagcttcaagtgtcagtttctgtcagtgagtgagcactgagacatttcaag 2249  
2070 TTTTGAATAGTTTGCAGTTGTGAGTTTGTTCATGAGGTGGCATCTGTGATCATTTCAAG 2129  
2250 cattgagacagactgacttcttctgtcttcttctccttctgatalataccttcaaacaa 2309  
2130 CATTTGAGGACAGACTGATCTCTCTTCTGTCTCTCTCTCGAATATACCTTCAACACAA 2189  
210 aatgtctatgaagaacactcaacctattcccaatctcagagagaacatgtctctatgtc 2369  
2190 AATGTCTATGAAGACACACTCACTCATCTATCCATTTCTCAGGAGAAACTGTCTTCATGTC 2249  
2370 gatggaataaccaggtctatgattctgggtgtccacaactcagacttctggaacagag 2429  
2250 GATGGAATAACCCAGGTATGAGATTCTGGGGTCCACAACTCAGACTTTGGAAACAAGG 2309  
2430 catgaccgcttactgaaggtttcttgaatgtgacaagaacacatggttattatcagagga 2489  
2310 CATGACCGCTTACTGAAAGTTTCTAGTTGTGACAAAGAACTGATATATATACGAGGA 2369  
2490 cagttatgaagaataattcagacatactgtctgagtaaaacaatgtcaatgtgaaccaagaag 2549  
2370 CAGTTATGAAGATATTTCACACTACTGTGAGTAAAAAACAAATGCCCTTAAACCAAGAAG 2429  
2550 ctctcccaagaattcagaagaacctagacactaggaacaaagaatlttatccacccca 2607  
2430 CTCTCCCAAGATTCAAGACACCTTACACTAGGCAAAAGCAATTTATGCCACACCA 2487

RESULT 13  
LOCUS I31901 9009 bp DNA PAT 06-FEB-1997  
DEFINITION Sequence 3 from patent US 5583209.  
ACCESSION I31901  
VERSION I31901.1 GI:1822692  
KEYWORDS

SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 9009)  
AUTHORS Lollar, V.S. and Runge, M.S.  
TITLE Hybrid human/porcine factor VIII  
JOURNAL Patent: US 5583209-A 3 10-DEC-1996;  
FEATURES location/Qualifiers  
source 1..9009  
BASE COUNT 2853 a 1907 c 1844 g 2405 t  
ORIGIN  
Query Match 49.0%; Score 2336.4; DB 10; Length 9009;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2337; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 270 catgcaatagagcttcacacgtcttcttctgtgcttcttgagattcgtcttagtgc 329  
DB 150 CATGCAATAGAGCTCTCCACCTGCTCTTCTGTGCTTTGGATTCTGCTTATGTC 209  
QY 330 caccagaagaatacctctgggtgagctgacatgtcaatggagactatgtcaaaatgtatct 389  
DB 210 CACCAGAAGATFACTACTCGGTGCACTGTAAGTGTATGAGGACTATATGCAAAAGTATCT 269  
QY 390 cgggtgagctgctgtgagcgaagattcctctcagagtgccaaatccttccattcaa 449  
DB 270 CGGTGAGCTCCTGTGAGCCGAAAGATTTCTCTCAGAGTGCAGAAATCTTTCCATTCAA 329  
QY 450 cactcagctgtgtaacaagaacactgtttgtagaattcaacagttcaaccttcaacat 509  
DB 330 CACTCAGTCTGTACCAAAAAGACTCTGTTGTGAAATTCACGTTTACCTTTTCAACAT 389  
QY 510 cgtctaaagcgaagccacccctgagatgagctcgtcagatgctcctacacatccagcctgaagttta 569  
DB 390 CGCTAAAGCCAAAGGCCACCTGAGATGGTGTGCTAGTGTCTACATCCAGGCTGAGTTTA 449  
QY 570 tgatacagtggtcaattacacttaagaacatggtctcccaatcctgtcagcttctatgtc 629  
DB 450 TGATACAGTGTCTATTCATTTAAGAAATGAGTCCCATCTGTCTCATGCTGT 509  
QY 630 tggctgctcctacatggaagaacctctctgagagagctggaatgatgatgaagccagctcaag 689  
DB 510 TGGTGTATCTTACTGGAAGCTTCTGAGGAGCTGAATATGATGATGAGACCTCAAG 569  
QY 690 ggaagaagaatgatataaagtcttccctgtgtggaagcacaacataltctcggcaggtcct 749  
DB 570 GGAGAAAGAAAGATGATTAAGTCTTCCCTGGTGGAGGCCATACATATGTCTGGCAGGTCT 629  
QY 750 gaaagaatggtctcaatctgacctctgacacatgtgcttactactacatacttctca 809  
DB 630 GAAAGAAATATGCTCAATGTGCTGTGACCCACCTGTGCTTACCTACATCATCTTTCTCA 689  
QY 810 tctggaacctgttaaaagacttgaattcagaagccatctgagagccctactagatgataaga 869  
DB 690 TGTGACCTGTGTAAGAACTGTGATTTAGGCTCTATTTGAGCCCTACTAGATGTAAGA 749  
QY 870 agggagctctgccaaggaagaacacagacctgtgacacaaattataactacttctgtct 929  
DB 750 AGGAGTCTGGCCAAAGAAAGACAGACGCTTGCACAAATTTATACTACTTTTGTCTGT 809  
QY 930 atttgatgaaggaagaagttggaacatcagaacaaagaactccttgatcagaatagga 989  
DB 810 ATTTGATGAAGGAGAAAGTTTGGCACTCAGAAAGAAAGAACTCTTGTGAGGATAGGA 869  
QY 990 tgcctgcatctgctcgggcccgtgaccttaaaatgcaacacagtaatggttatgttaaagcgtc 1049  
DB 870 TGCTGCAATCTGCTCGGGCTTGAATATGACACAGTAAATGTTATTAACAGGTC 929  
QY 1050 tctgcaagctctgattggaatgacacaggaatcagttcatcttgacatgtgaatgag 1109  
DB 930 TCTGCCAGGTCTGTGATGATGCCACAGGAATCAGTCTATTTGGCATGTGATTTGGAATGGG 989

|    |      |  |      |
|----|------|--|------|
| OY | 1110 | caacaccccttgaaigtgcactcaataatccctcgaaagtcacacattctctgtgaggaacca  | 1169 |
| Db | 990  | caacaccccttgaaigtgcactcaatattccctcgaaagtcacacattctctgtgaggaacca  | 1049 |
| OY | 1170 | tcgcaagagctctcttgaaatctccgccaataactcttctaactgctctcaaacacttgat    | 1229 |
| Db | 1050 | tcgccaagagctctcttgaaatctccgccaataactcttctaactgctctcaaacacttgat   | 1109 |
| OY | 1230 | ggaccttggacagtlcttaactglttltgcalaactctcccaaccaaatgataggcatlga    | 1289 |
| Db | 1110 | ggaccttggacactttctactgtttttgtcattcttccacccaacatgatgatgca         | 1169 |
| OY | 1290 | agcttatgtlcaaaagttagacagctgtccagagggaacccccaaactaagaatgaaataatga | 1349 |
| Db | 1170 | accttatgttcaaaagttagacagctgtccagagggaacccccaaactaagaatgaaataatga | 1229 |
| OY | 1350 | agaagcggagaactatgatatgatcttaactgattctcgaatggatgtgtgcaggtlga      | 1409 |
| Db | 1230 | agaagcggagaactatgatatgatcttaactgattctcgaatggatgtgtgcaggttga      | 1289 |
| OY | 1410 | tgatgacaactctctcctctcttaaccaaltgcgtcagttgcgaagaacatccaaac        | 1469 |
| Db | 1290 | tgatgacaactctctcctctcttaaccaaltgcgtcagttgcgaagaacatccaaac        | 1349 |
| OY | 1470 | ttgggtgacatataattgcttccaaaggagagatggagactatgtcccttaattgcttgc     | 1409 |
| Db | 1350 | ttgggtgacatataattgcttccaaaggagagatggagactatgtcccttaattgcttgc     | 1409 |
| OY | 1530 | ccccgaatgcagaaagtataaagaatcaatattgaacatggccctcagcagatctgtag      | 1589 |
| Db | 1410 | ccccgaatgcagaaagtataaagaatcaatattgaacatggccctcagcagatctgtag      | 1469 |
| OY | 1590 | gaagtaacaaaagatccgatttatggcatacacagatgaaacctttaagatcgtlgaagc     | 1649 |
| Db | 1470 | gaagtaacaaaagatccgatttatggcatacacagatgaaacctttaagatcgtlgaagc     | 1529 |
| OY | 1650 | tattcagcatgaaatcaggaatctcttggaccttaactttaagggaagtcttgagacacact   | 1709 |
| Db | 1530 | tattcagcatgaaatcaggaatctcttggaccttaactttaagggaagtcttgagacacact   | 1589 |
| OY | 1710 | gttgatcatactttaagaaatcaagcaagcagaacatatcaactctacccctcagaaatcac   | 1769 |
| Db | 1590 | gttgatcatactttaagaaatcaagcaagcagaacatatcaactctacccctcagaaatcac   | 1649 |
| OY | 1770 | tgatgtcgtctccttgtlatcgaaggagatlaacaaagtgtaaacatttgaagatct        | 1829 |
| Db | 1650 | tgatgtcgtctccttgtlatcgaaggagatlaacaaagtgtaaacatttgaagatct        | 1709 |
| OY | 1830 | tccaattctgcgcaggagaaatctcaaaataaattgaacagtgatctlaagaatgggcc      | 1889 |
| Db | 1710 | tccaattctgcgcaggagaaatctcaaaataaattgaacagtgatctlaagaatgggcc      | 1769 |
| OY | 1890 | aactcaaatagatctctcgtgtgccttgaccgcgctatlaactcagttcgtltaataggagag  | 1949 |
| Db | 1770 | aactcaaatagatctctcgtgtgccttgaccgcgctatlaactcagttcgtltaataggagag  | 1829 |
| OY | 1950 | agatctagcttcaaggactaattggtccctctctcaactctgtcacaagaatctgtagatca   | 2009 |
| Db | 1830 | agatctagcttcaaggactaattggtccctctctcaactctgtcacaagaatctgtagatca   | 1889 |
| OY | 2010 | aagaggaacacagataatgttcaacaagaaggaatgtacatcgttcttctgtatttgatga    | 2069 |
| Db | 1890 | aagaggaacacagataatgttcaacaagaaggaatgtacatcgttcttctgtatttgatga    | 1949 |
| OY | 2070 | gaacccaagctggtaacctcaacagagataatacagcgttctccccaactcagcttgagt     | 2129 |
| Db | 1950 | gaacccaagctggtaacctcaacagagataatacagcgttctccccaactcagcttgagt     | 2009 |
| OY | 2130 | gcagcttgaagatcagaagttcccaagcttccaaatcaatgacagaaatcaatggctatgt    | 2189 |
| Db | 2010 | gcagcttgaagatcagaagttcccaagcttccaaatcaatgacagaaatcaatggctatgt    | 2069 |

[illegible]





| ORGANISM     | REFERENCE           | AUTHORS  | TITLE   | JOURNAL     | MEDLINE  | COMMENT |
|--------------|---------------------|--|---|-------------|----------|---------|
| Homo sapiens | 1 (bases 1 to 9009) | Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. | Currently hemophiliacs are treated with factor VIII:C purified from blood. The human factor VIII:C cDNA has now been isolated, cloned and expressed. The gene was also isolated. It consists of a large number of exons spanning over 180 Kb. The authors kindly supplied the sequence directly to Genbank by computer network. | Factor VIII | 85061550 |         |
| gene         | 151. .7206          |  | 151. .7206  |             |          |         |
| CDS          | 151. .7206          |  | 151. .7206  |             |          |         |
| FEATURES     | 1. .9009            |  | 1. .9009  |             |          |         |
| source       | 1. .9009            |  | 1. .9009  |             |          |         |
| gene         | 151. .7206          |  | 151. .7206  |             |          |         |
| CDS          | 151. .7206          |  | 151. .7206  |             |          |         |
| FEATURES     | 1. .9009            |  | 1. .9009  |             |          |         |
| source       | 1. .9009            |  | 1. .9009  |             |          |         |



QY 1290 agctatgtcaaaagtagacaagctgtccagaagaaacccaactacgaatgaanaaataatga 1349  
|||||  
Db 1170 AGCTTATGTCTCAAACTAGACAGCTCTCCAGAGAACCCCACTAGCATGAAAAATATATGA 1229  
QY 1350 agaagcggaaagactatgatgatgacttactgacttctgaaatggaatggtcgaagtttga 1409  
|||||  
Db 1230 AGAAGCGGAAGACTATGATGATGATCTTACTGATTCGAAATGATGTGTCAGCTTTGA 1289  
QY 1410 tgatgaacactctctcccttaccnaatctgctcagtttgccaagaagatcctaataac 1469  
|||||  
Db 1290 TGATGACAACTCTCTCTCTTATCCAAATTCGCTCAGTTGCCAAGACATCTCTAAAC 1349  
QY 1470 ttgggtacattacattgtcgtgaaaggaaggaaggaaggaaggaaggaaggaaggaaggaag 1529  
|||||  
Db 1350 TTGGGTACATTACATTACATTACATTACATTACATTACATTACATTACATTACATTAC 1409  
QY 1530 ccccgatgacagaagatataaagltcaalattgnaacatgycctcagcggatltgtag 1589  
|||||  
Db 1410 CCCCGATGACAGAGATTATTAAGTCAATATTTGAACAAATGGCCCTCAGCGGATTTGATG 1469  
QY 1590 gaagtaacaaaagatccgaatttatgacacacagatgaacactttaagactgtgaagc 1649  
|||||  
Db 1470 GAACTACAAAAAAGTCCGATTATGCGATACACAGATGAACCTTTAAGACTCCTGAAGC 1529  
QY 1650 tattcagcatgaatcagaagactcttggaaccttacttattggggaagttggaagacacact 1709  
|||||  
Db 1530 TATTACAGCATGATACAGGAATCTTGGACCTTTACTTTATGGGAAATTTGAGACACACT 1589  
QY 1710 gtgatataattaaagaatcaagaagaacacalataacatctacacctcaaggaatcac 1769  
|||||  
Db 1590 GTTGATTATATTATTAAGATCAAGACAGACACATATTAACATTAACCTCAGCGAATCAC 1649  
QY 1770 tgaatgcgtcccttctgataccaaggaatataccaaaagtgtaaacatttgaagaattc 1829  
|||||  
Db 1650 TGATGTCGGTCTTTGATTCAAGAGATTACCAAAGGTGAACATTTGAAGGATTT 1709  
QY 1830 tccaattctccaggaagaataatcaataataatgaagtgactgtagaagaatggcc 1889  
|||||  
Db 1710 TCCAAATCTGCGAGAGAAATATTCAATATTAATGACAGTGTAGAGATGGGCC 1769  
QY 1890 aactaaatcagatcctcgtgctcgaaccgctactactagttcgtltaatatgagaag 1949  
|||||  
Db 1770 AACTAAATCAGATCCTGGGCGCTGACCGGCTATTACTAGTTGCTTAATATGAGAG 1829  
QY 1950 agatctagctcagaagactcaattggcctctcccaatctgctacaagaatcgtlagatca 2009  
|||||  
Db 1830 AGATCTAGCTTCAGGACTCATTTGGCCCTCTCTCATCTGCTACAAAGAAATCTGTAGATCA 1889  
QY 2010 aaagggaaacaaataatgtcagaagaagaagatgcatcctgttctgtatttga 2069  
|||||  
Db 1890 AAGAGAAACCAAGATTAATGTCAAGCAAGAGAAATGTATCTCTGTTCTGTATTGATGA 1949  
QY 2070 gaacccaagctgtgtacctcacaagaatatataaagcttctcccaatccaagctgaag 2129  
|||||  
Db 1950 GAACCGAAGCTGTGACTCTCAGAGAAATATACAACGCTTTCTCCCAATCCAGCTGAGT 2009  
QY 2130 gcagcttgaggaatccagaagtccaagcctcaacatcatgcaagaatcaatggctatgt 2189  
|||||  
Db 2010 GCAGCTTGAGGATTCAGAGTTCCAAAGCCTCAACATCATGACAGCATCAATGCTATGT 2069  
QY 2190 tttagaatgttcgaagtgtcagttgttggcaggaagtggaactggtacacttgaag 2249  
|||||  
Db 2070 TTTTGAATGTTTGCAAGTTTGCAAGTTTGTGATGAGGATGAGGATGATGATCTTAAG 2129  
QY 2250 cattggaacacagactgacttcttctgcttcttctctgatalataacttcaacacaa 2309  
|||||  
Db 2130 CATTGAGACACAGACTGATCTCTTCTGCTCTCTCTGATATATCTTCAACACAA 2189  
QY 2310 aatggtctatgaagaacacactcaacctatcccaattctcaaggagaactgtctcatgtc 2369  
|||||  
Db 2190 AATGTCTATGAAGACACTCAACCTATTCCTCCTCAGGAGAAACGTCTTCATGTCT 2249

QY 2270 gatgaaaacccaagctctatgaltctgggtgcccacaactcagacttccgaaacagag 2429  
|||||  
Db 2250 GATGAAAACCCAGCTGTATGATTTCTGGGTTGCCACAACTCACTTTGGAAACAGAG 2309  
QY 2430 catgacgccttactgaagtttctagttgtgacaagaacacatggtatattagagga 2489  
|||||  
Db 2310 CATGACCGCTTACTGAAAGTTTCTAGTTGTGACAAAGACACTGTGATTTATCGAGGA 2369  
QY 2490 cagttatgaagataattcagatactgctgagatgaanaacatgccaatgaacaaag 2549  
|||||  
Db 2370 CAGTTATGAAGATATTTCAGCATCTTGCTGAGTAAAAACAATGCCATTTGAACCAAGAG 2429  
QY 2550 ctctcccaagatccaagacacccctagcaactaggaacaaagaatttaatgaccacca 2607  
|||||  
Db 2430 CTCTCCAGATTCAGAGACACCTAGCAGCTAGGCAAAAGCAATTTAATGCAACACA 2487

Search completed: November 18, 2001, 01:36:22  
Job time: 30124 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2001, 04:50:04 ; Search time 389.82 seconds  
(without alignments)  
2566.193 Million cell updates/sec

Title: US-09-689-430-1\_COPY\_419\_4835

Perfect score: 4417  
1 catgcaatagagctctcca.....aggcacagcactactga 4417

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA:\*

1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/Backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysts of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | ID | Description                             |
|------------|--------|-------------|--------|----|---|
| ✓ 1        | 4412.8 | 99.9        | 4629   | 2  | US-08-484-891-7<br>Sequence 7, Appli    |
| ✓ 2        | 4407.4 | 99.8        | 9354   | 1  | US-08-683-839B-2<br>Sequence 2, Appli   |
| ✓ 3        | 4357.4 | 98.7        | 4670   | 3  | US-08-717-294-41<br>Sequence 41, Appli  |
| ✓ 4        | 4319.8 | 97.8        | 4999   | 4  | US-09-470-618-14<br>Sequence 14, Appli  |
| ✓ 5        | 4319.8 | 97.8        | 4999   | 4  | US-09-364-862-14<br>Sequence 14, Appli  |
| 6          | 4249.4 | 96.2        | 11933  | 4  | US-09-470-618-13<br>Sequence 13, Appli  |
| 7          | 4249.4 | 96.2        | 11933  | 4  | US-09-364-862-13<br>Sequence 13, Appli  |
| 8          | 3743.8 | 84.8        | 5035   | 2  | US-08-882-083-1<br>Sequence 1, Appli    |
| 9          | 3743.8 | 84.8        | 5035   | 2  | US-08-558-107-1<br>Sequence 1, Appli    |
| 10         | 3743.8 | 84.8        | 5035   | 3  | US-09-243-539-1<br>Sequence 38, Appli   |
| 11         | 3189.8 | 72.2        | 4334   | 2  | US-08-670-707A-38<br>Sequence 38, Appli |
| 12         | 3189.8 | 72.2        | 4334   | 4  | US-09-037-601-38<br>Sequence 38, Appli  |
| 13         | 2603.8 | 58.9        | 4451   | 3  | US-08-717-294-42<br>Sequence 42, Appli  |
| 14         | 2336.4 | 52.9        | 9009   | 1  | US-07-864-004B-3<br>Sequence 3, Appli   |
| 15         | 2336.4 | 52.9        | 9009   | 1  | US-08-212-133A-3<br>Sequence 3, Appli   |
| 16         | 2336.4 | 52.9        | 9009   | 1  | US-08-212-133A-1<br>Sequence 1, Appli   |
| 17         | 2336.4 | 52.9        | 9009   | 1  | US-08-474-503-1<br>Sequence 1, Appli    |
| 18         | 2336.4 | 52.9        | 9009   | 2  | US-08-670-707A-1<br>Sequence 1, Appli   |
| 19         | 2336.4 | 52.9        | 9009   | 4  | US-09-037-601-1<br>Sequence 1, Appli    |
| 20         | 2336.4 | 52.9        | 9009   | 5  | PCT-US93-03275-3<br>Sequence 3, Appli   |
| 21         | 2336.4 | 52.9        | 9009   | 5  | PCT-US94-13200-1<br>Sequence 1, Appli   |
| 22         | 2336.4 | 52.9        | 7056   | 1  | US-08-121-202-1<br>Sequence 1, Appli    |
| 23         | 2334.8 | 52.9        | 8241   | 6  | 5171844-1<br>Patent No. 5171844         |
| 24         | 2333.2 | 52.8        | 8967   | 1  | US-08-366-851A-1<br>Sequence 1, Appli   |
| 25         | 2275.2 | 51.5        | 6999   | 1  | US-08-276-584A-1<br>Sequence 1, Appli   |
| 26         | 1848   | 41.8        | 7032   | 4  | US-09-324-867-1<br>Sequence 1, Appli    |
| 27         | 1801.8 | 40.8        | 7493   | 1  | US-08-212-133A-7<br>Sequence 7, Appli   |

|    |        |      |      |   |   |
|----|--------|------|------|---|---|
| 28 | 1801.8 | 40.8 | 7493 | 1 | US-08-474-503-5<br>Sequence 5, Appli    |
| 29 | 1801.8 | 40.8 | 7493 | 2 | US-08-670-707A-5<br>Sequence 5, Appli   |
| 30 | 1801.8 | 40.8 | 7493 | 4 | US-09-037-601-5<br>Sequence 5, Appli    |
| 31 | 1801.8 | 40.8 | 7493 | 5 | PCT-US94-13200-5<br>Sequence 5, Appli   |
| 32 | 1751.2 | 39.6 | 6402 | 2 | US-08-670-707A-36<br>Sequence 36, Appli |
| 33 | 1751.2 | 39.6 | 6402 | 4 | US-09-037-601-36<br>Sequence 36, Appli  |
| 34 | 1240.4 | 28.1 | 1623 | 1 | US-08-121-202-3<br>Sequence 1, Appli    |
| 35 | 893.4  | 20.2 | 1130 | 1 | US-07-864-004B-1<br>Sequence 1, Appli   |
| 36 | 893.4  | 20.2 | 1130 | 1 | US-08-251-937A-1<br>Sequence 1, Appli   |
| 37 | 893.4  | 20.2 | 1130 | 1 | US-08-212-133A-5<br>Sequence 5, Appli   |
| 38 | 893.4  | 20.2 | 1130 | 1 | US-08-474-503-3<br>Sequence 3, Appli    |
| 39 | 893.4  | 20.2 | 1130 | 2 | US-08-670-707A-3<br>Sequence 3, Appli   |
| 40 | 893.4  | 20.2 | 1130 | 4 | US-09-037-601-3<br>Sequence 3, Appli    |
| 41 | 893.4  | 20.2 | 1130 | 5 | PCT-US93-03275-1<br>Sequence 1, Appli   |
| 42 | 893.4  | 20.2 | 1130 | 5 | PCT-US94-13200-3<br>Sequence 3, Appli   |
| 43 | 301.8  | 6.8  | 6909 | 2 | US-08-804-136-1<br>Sequence 1, Appli    |
| 44 | 301.8  | 6.8  | 6909 | 2 | US-08-658-340-1<br>Sequence 1, Appli    |
| 45 | 301.8  | 6.8  | 6909 | 3 | US-08-746-111-26<br>Sequence 26, Appli  |

## ALIGNMENTS

RESULT 1  
US-08-484-891-7  
Sequence 7, Application US/08484891  
Patent No. 5935935  
GENERAL INFORMATION:  
APPLICANT: Connolly, Sheila  
APPLICANT: Kaleko, Michael  
APPLICANT: Smith, Theodore  
TITLE OF INVENTION: Adenoviral Vectors for  
TITLE OF INVENTION: Treatment of Hemophilia  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Carella, Byrne, Bain, Gilfillan,  
ADDRESSEE: Cecchi, Stewart & Olstein  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,891  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/218,335  
FILING DATE: 25-MAR-1994  
APPLICATION NUMBER: 08/074,920  
FILING DATE: 10-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Olstein, Elliot M.  
REGISTRATION NUMBER: 24,025  
REFERENCE/DOCKET NUMBER: 271010-273  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4629 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
MOLECULE TYPE: linear  
FEATURE:  
NAME/KEY: Factor VIII CDNA with  
NAME/KEY: B domain deleted

US-08-484-891-7

Query Match 99.9%; Score 4412.8; DB 2; Length 4629;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 atgcaaatgagctccacacgctcttctctgtgcttcttgagatctcgttaagtc 61  
DB 1 ATGCAAAATGAGCTCTCCACCTCTCTCTGCTGCTTTGGCATTTCTCTTAGTGGCC 60  
QY 62 accaagaatactaccctgggtgagtgaaactgcaatggactatagcaaatgactc 121  
DB 61 ACAAGAGATGATCTACTGCTGGTGGAGTGGAATGTGATGGGACTTATGCCAAGTGATCTC 120  
QY 122 ggtgagctgctgtgagcgaagaattcctcctagagtgccaaatctttccatc 181  
DB 121 GGTGAGCTGCTGTGGAGCGCAAGATTTCTCTAGAGTGCCAAATCTTTTCCATTCAC 180  
QY 182 accctagctgtgtacaaaagaactctgtgtgtagaatcaggttcaaccttcaac 241  
DB 181 ACCCTAGCTGTGTACAAAAGACTCTGTGTGTAATTCAGGTTCCACTTTTTCACATC 240  
QY 242 gctaaagccaaagccaccctggatgggtctgtgctaggtctcaaccacagcgtgaggttat 301  
DB 241 GCTAAGCCAAAGCCACCTGGATGGGTCTGTAGGTCTTACCATCCAGGCTTGAGGTTTAT 300  
QY 302 gatacagtggtcaatcaacttaagaacalggcttccatctgttcaagttctca 361  
DB 301 GATACAGTGTGATACACTTAAAGAACATGGCTTCCATCTGTCAAGTCTTCAATGCTGT 360  
QY 362 ggtatctactctggaagaactcttgaggagcgtgaatagtatgaaccagtc 421  
DB 361 GGTGATCTCTACTGTGAACCTTCTGAGGAGCTGGAATATGATGATGACCACTCAAGG 420  
QY 422 gagaagaagaatgataaagctcttccctgggtggaagccatacatatcttgc 481  
DB 421 GAGAAAGAAGATGATTAAGTCTTCCCTGCTGGAGAGCCATATATGTCTGGAGTCTG 480  
QY 482 aaagaagaatggtccaaatggcctctgacccactgtgcttcaactcaatctt 541  
DB 481 AAGAGAAATGCTCAATGCTCTGTGACCCACTGTGCTTACCTATCATATCTTCTGCAT 540  
QY 542 gtggaccctgtgtaaaagacttgaatcagcctcatatggagccctactagatg 601  
DB 541 GTGGACCTGTGTAAGACTTGAATTCAGGCTTATGGAGCCCTACTACTATGTAGAA 600  
QY 602 gggagcttggccaagaaaagacacagaccttgacacaaatltactacttttgc 661  
DB 601 GGGAGTCTGGCCAAAGAAAGACAGACCTTGCACAAATTTTACTACTTTTGTCTGTA 660  
QY 662 ttgtatgaagggaagaatgtgcaactcagaacaagaactccttgaatgagga 721  
DB 661 TTTGTATGAAGGAAAAATGTGGCAGCTCAGAAACAAAGAACTCTTGTATGACAGATAGGAT 720  
QY 722 gctgatatctgctcgggcccgtgcttaaaatgacacacagtaatggttatg 781  
DB 721 GCTGATCTGTCTCGGGCCCTGCTTAAATATGACACAGTCAATGGTATGTAAACAGGTCT 780  
QY 782 ctgcaagctctgattgagtgccacaggaatcagctcatatgtgcaatgga 841  
DB 781 CTGCAAGCTCTGATTTGATGCCACAGGAATATGATTTGGCATGTGATTGGAATGGG 840  
QY 842 accactcctgaagtgcaactcaatctcctcgaagtgcaacatcttctgtga 901  
DB 841 ACCACTCTGAAGTGACACTCATATTTCTCGAAGGTCAACATTTCTTGAGGAACAT 900  
QY 902 cgcaagagctcttggaatatctgcgaataaacttctccttaagtcataaact 961  
DB 901 CGCAGAGGCTCTTGGAATCTCGCAATTAATTTCTTACGTCTCAAAACACTCTTGAG 960  
QY 962 gaccttgacagattctactgtttgtgatactctctccacaacaatgata 1021  
DB 1021 GACCTTGACAGATTCTTACTGTTTTGTGATATCTCTTCCACCACACATGATGATGAA 1020

QY 1022 gctatgtcaaaatgagcaagctgtccagaggaaccccaactacgaaatga 1081  
DB 1021 GCTATGTCAAAATGAGACACTGTCCAGAGAACCCCACTACGATGAAAAATATGAA 1080  
QY 1082 gaagcgggaagactatgatactgatacttctgaatggaatgagtgaggtt 1141  
DB 1081 GAAGCGGAAGACTATGATGATGATCTTACTGATTCGAATGATGATGTGTCAGGTTGAT 1140  
QY 1142 gatgaaactctctcctcttaccatcgaatcgctgaggttggcaagaagatc 1201  
DB 1141 GATGACACTCTCTCTCTTATTCAAATTCGCTCAAGTTCAGAGATCTTAAACT 1200  
QY 1202 tgggtatattacatctgtctgtaagagaggaacttggagatgctcctc 1261  
DB 1201 TGGGTATATTACATTTGTGCTGAAGAGAGACTGGGACTATGCTCTTACTCCCTGCC 1260  
QY 1262 ccgagatgacagaagtataaagatcaatattgaaacatggccctcagcgat 1321  
DB 1261 CCGAGTACAGAAAGTTATAAAGTCAATATTGGAACAAATGGCCCTCAGCGATTGGTAGG 1320  
QY 1322 aagttacaaaagctcgaattatggtatgacacagatgaaaccttaagaact 1381  
DB 1321 AAGTACAAAAGTCCGATTATGCGATACAGATGAACCTTTAGACTGTGAAGCT 1380  
QY 1382 attcagatgaaatcagaagctcttggagcttcaacttgaagggaggttga 1441  
DB 1381 ATTACAGATTAATCAGGAATCTTGGAGCTTACTTTATGGGAGTTGACACACTG 1440  
QY 1442 ttgatatatttaagaatcaagcagaacacataaactcaactcaactc 1501  
DB 1441 TTGATTATATTAAAGATCAAGCAACACCATATAATCATCTACCTCAGCAATCACT 1500  
QY 1502 gatgtccgctcttctgattcaagaagataccaagaaggtgtaaaacatttga 1561  
DB 1501 GATGTCCGCTCTTGTATTCAGAGATTAACCAAAAGGTGTAACATTGTAAGGATTTT 1560  
QY 1562 ccaattctgcagagagaatatccaataataatgagcagtgactgagaatg 1621  
DB 1561 CCATTTCTGCAGAGAAATATTAATTAATGAGACAGTCACTGTAAGATGGGCA 1620  
QY 1622 actaaatcagatctcctggtgcttgacccgtatctactagtttgcgtta 1681  
DB 1621 ACTAAATCAGATCTCTGGGCCCTGAGCCGCTATTACTGATTTGTTGTAATGAGAGA 1680  
QY 1682 gatcagcttcaagactcaatggccctcccatctgttaaaagaactctgata 1741  
DB 1681 GATCTAGCTTCAAGACTCATTTGGCCCTCTCTCATCTGTACAAAGAAATCTGTAGATCA 1740  
QY 1742 agagaaaccagataatgtcagacaagaaggaatgtcaatctgtttctgt 1801  
DB 1741 AGAGAAACAGATATATGACAGAAAGAGAAATGATCATCTGTTTCTGATTTGATGAG 1800  
QY 1802 aaccgaagctgttactcaacagaataataaagcttctcccaatccagct 1861  
DB 1801 AACCGAAGCTGTACTCTCAGAGAAATATACAGCGTTTCTCCCAATCCAGCTGAGAG 1860  
QY 1862 cagcttgagatccagaagttccaaagcttcaacatcatgacagataatg 1921  
DB 1861 CAGCTTGAGGAATCCAGAGTTCAGAGCTCAACATCATATGACAGATCAATGCGATGAT 1920  
QY 1922 ttgataagtttgcaagttgtcagttgtttgtgaatgagtgcaatctgata 1981  
DB 1921 TTTGATAGTTTGCAAGTTTGCAATTTGTTGATGAGTGCGATCTGTGATTTTAAAC 1980  
QY 1982 attgagacagactgaaactcttctctgtctctctcctgatalaacttca 2041  
DB 1981 ATTGAGACACAGACTGACTCTTCTGCTCTCTCTGATATACCTCAAAACACAAA 2040  
QY 2042 atggtctatgaagaacaataccatctccatctcaagagaacagcttca 2101  
DB 2041 ATGCTTATGAAGACACACTCAACCTATTTCCATTTCTCAGAGAAAGTCTTCAATGTCG 2100

QY 2102 atggaaccacagctctatgatctctg99gtgcacaaactcagaacttgc9aaacagagc 2161  
|||||  
Db 2101 ATGGAACCCAGCTCTATGATTTCTGGGTGCCACACTCGACTTTTCGAAACAGAGC 2160  
QY 2162 atgacgccttactgaaggttctctagttgtgaacaaacacatggtgatatactacagagc 2221  
|||||  
Db 2161 ATGACCGCCTTACTGAAAGGTTTCTAGTGTGACAGAGACACTGGTGAATTAATTCAGAGGAC 2220  
QY 2222 agttatgaagatacttcagacatacttgcgtgagttaaaacaaatgcattgaacaaagc 2281  
|||||  
Db 2221 AGTTAAGAAGATATTTACAGATCTTGTGAGTAAACAAATCCATTCGAAACCAAGC 2280  
QY 2282 tctcccaaatccaagaaacctagcactaggcaaaagaatttaagccaccac 2341  
|||||  
Db 2281 TTCTCCAGATTCAAGACACCTAGACAGCAAAAGCAATTAATGACCCACCA 2340  
QY 2342 gtcttgaacgcacataacg9gaataaactcgtactactctcagtcagaatcaagagaa 2401  
|||||  
Db 2341 GTCTTGAACGCCATCAACGGGAAATACCTGACTTCTTACGTACAGATCAAGAGAA 2400  
QY 2402 attgactatgatataccatcatalcagltgaagaagaagaatttgacattatgat 2461  
|||||  
Db 2401 ATGACTATGATGATACCATATCATGAGTGAATGAAGAGAGATTTTGACATTTATGAT 2460  
QY 2462 gagagataaatcagaagcccccagcttccaagaagaacagacatacttatgct 2521  
|||||  
Db 2461 GAGGATGAATAATCAGAGCCCCGACCTTTCAAAAGAAAACAGACATATTTATTTGCT 2520  
QY 2522 gacgtgagaagcgtctg9gataatgagatgagctccacatgttctaagaacag 2581  
|||||  
Db 2521 GCAGTGAAGAGCTCTGGGATTAATGGAGATGAGTCCCAATGTTCTCAAGAAACAG 2580  
QY 2582 gctcagaatgagcagtgctccctcagttcaagaagtgcttccagaatttactgaatg 2641  
|||||  
Db 2581 GCACAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTTTCAGAGAAATTTACTGATGC 2640  
QY 2642 tcccttactcagcccttatacgttgaagaactaaatgaacatttggacccctgggcca 2701  
|||||  
Db 2641 TCCCTTACTCAGCCCTTATATCCGTGAGACTTAATGAACATTGGAGCTCCGGGGCCA 2700  
QY 2702 tatataagacagaagtctgaagataatcaltcagtaacttccagaatcagcctcgt 2761  
|||||  
Db 2701 TATATAGAGCAAAATTTGAAGTATATCATGTAACTTTCAGAAATCAGGCTCTCCT 2760  
QY 2762 cccattcctctatctcagccttatttcttgaagaagatcagaagcagaagacaga 2821  
|||||  
Db 2761 CCTATTCCCTTATCTAGCCTTATTTCTTATGAGGAAGATCAGAGCAGAGAGAA 2820  
QY 2822 cctagaaaaaacttctcagaagcctaatagaacaaacttacttcttgaagaatgacaat 2881  
|||||  
Db 2821 CCTAGAAAAAACTTTTCTCAAGCCTAATGAACCAAAACCTTACTTTTGAAAGTGCAACAT 2880  
QY 2882 cataatgccccactaaagatgagttgactgaagcctgaagcctgagcttatctctgagt 2941  
|||||  
Db 2881 CATATGGCACCCCTAAAGATGAGTTGACTGCAAAAGCCTGGGCTTATTTCTGTGATGT 2940  
QY 2942 gaccttgaaaaaagatgtgacatcagaagcctgatatgaaacccctctggtctgccaactaac 3001  
|||||  
Db 2941 GACCTGAAAAAAGATGTGACATCAGGCTGATGAGACCCTTCTGTGTCACACATAAC 3000  
QY 3002 acaatgaacccctgctcactatg9gagacaaagtacaggaatttgcctgttcttcacc 3061  
|||||  
Db 3001 ACACCTAACCCTGCTCATGGAGACAAAGTACAGTACAGGAATTTCTCTGTTTTTCACC 3060  
QY 3062 atcttgaatgaagcacaagcctgttacttcaactgaaataatgaaagaacatgcagagct 3121  
|||||  
Db 3061 ATCTTTGATGAGACCAAAAGCTGTACTTCACGAAAATTAATGAAAGAAACTCAGAGGCT 3120  
QY 3122 cccctgaataatcagaatggaagatcccaactttaaagaagatatgccttcaatgcaatc 3181  
|||||  
Db 3121 CCTGCAATATCATCAGATGGAAGATCCCACTTTTAAAGAGAAATTAATGCTTCATGCAATC 3180

QY 3182 aatgctacataatggaatacactacctgcttagtaatggtctcagaatcaaaagattcga 3241  
|||||  
Db 3181 AATGGCTACTAATATGATATACACTACCTGCGCTTAGTAATGGCTCAGATCAAGAGATTGCA 3240  
QY 3242 tggatctgctcagcatlbggcagaacatgaacatccatctatctcaatcagltgagacat 3301  
|||||  
Db 3241 TGGTATCTGCTCAGCATGGGACAGCAATGAACATCCATCTATTCATTTCACTGAGCAT 3300  
QY 3302 gtgttcaactgtaagaaaaaagaggtatataaatlbgcactgtacaaatctctatccaggt 3361  
|||||  
Db 3301 GTGTTCACGTACGAAAAAAGAGAGATTAATAATGGCACTGTATCAATCTCTATCCAGGT 3360  
QY 3362 gtcttgaagcagltggaatggttccactcaagaagcctgaatttggcgggtggaatgctt 3421  
|||||  
Db 3361 GTTTTGAACAGTGGAAATGTTTACATCCAAAGCTGGAATTTGGCGGGTGGAAATGCTT 3420  
QY 3422 attg9cagatctataatgcttggatlgagacacttcttggltgtaacgaataagltgt 3481  
|||||  
Db 3421 ATTGGGAGCATCTACATGCTGGGATGAGCACACTTTTCTGGTGTACAGCAATTAAGTGT 3480  
QY 3482 cagaatcccttgggaatggtctctggaacattagagaatttcaagatcagcttcaaga 3541  
|||||  
Db 3481 CAGACTCCCTGGGAATGGCTTCTGACACATTAAGATTTTCAGATTACAGCTTCAGGA 3540  
QY 3542 caatlggaagatg9gcccgaagcctg9cagaacttcatattccggaataatcaatgcc 3601  
|||||  
Db 3541 CAATATGAGACAGTGGGCCCAAGAGCTGGCCAGACTTATTAATTCGGATTAATCAATGCC 3600  
QY 3602 tggagacccaagagcccttcttctgatatcaagltgagatcgttggacaaatgatatt 3661  
|||||  
Db 3601 TGGAGACCAAGAGGCCCTTTTCTTGATCAAGTGTGATCTGTTGGCAACAAATATTAAT 3660  
QY 3662 cacggaatcaagaacccaaggttgcctcagaaagltctcagactctacatctcagttt 3721  
|||||  
Db 3661 CACGGCATCAAGACCAAGGCTGCCGTCAGAAAGTTCACAGCCCTCAATCTCAAGTTT 3720  
QY 3722 atcataatgatatgcttcttgaatggaagaagtggcaagcattacgaagaattccacgga 3781  
|||||  
Db 3721 ATCATCTGTATATGTCTTGTATGGAGAAATGGCAGACTTATGAGGAATTTCCACTGGA 3780  
QY 3782 accatlaatgltcttcttcttggacaaatlgatcatcaltggaataaacaataatlttaac 3841  
|||||  
Db 3781 ACCTTAATGCTCTTCTTTGGCAATGTGATTCATCTGGATTAACCAATATTTTAAC 3840  
QY 3842 cctccaatatactcgtatcatalccgttgcacccaactcatatagatctgcagact 3901  
|||||  
Db 3841 CCTCAATTAATTTCTGATCATCTCCGTTTGCACCAACTCATTAATACATTCGACGACT 3900  
QY 3902 ctccgaatgagttgagtggtctgtaattaaatgattgagcagctgccaatgggaatggag 3961  
|||||  
Db 3901 CTTGGCATGAGTGTGATGGGCTGTGATTAATAATGTGACGATGCGATTGGGATGGAG 3960  
QY 3962 agtaaaagcaatcagaatgacagatcactgcttcaacttcaacttcaactatglttggc 4021  
|||||  
Db 3961 AGTAAGCAATATCAAGATGACAGATTAACGTTCATCTCACTTAACCATATATGTTGGC 4020  
QY 4022 acctggtctctccaagaagcctgcagctcacctccaaggaagaaatgcttggagact 4081  
|||||  
Db 4021 ACCTGTCTCTCTCAAAAGCTTCGACTTCACCTCCAAAGGAGGATATGCTGTGAGACT 4080  
QY 4082 caagtgaataatccaagaaggtggcttgcagatggacttccaagaagaataatgaagtaca 4141  
|||||  
Db 4081 CAGGTGAATATATCAAAAGAGTGTGCAAGTGAAGTCAAGAGCAATGAAGAGTACACA 4140  
QY 4142 g9gataactactcaaggaagaaatcctctcttcaacagcaatgtaatggaaggttctc 4201  
|||||  
Db 4141 GAGTAACTACTCTCAGGAGATGAATAATCTGCTTAACAGCAATGATGGAAGAGTCTCTC 4200  
QY 4202 atctcagaagcagatggcactcagltgagctctcttcttcaagaatggcaagtaag 4261  
|||||  
Db 4201 ATCTCCAGCAGTCAAGATGAGTGCATGAGTGCATCTCTTTTTCAGATGCAAGTAAAG 4260  
QY 4262 gttttcagaagaaatcaagaactccttcaacccgtgtgtgaaactctcaagcccaacgtt 4321

|||||  
Db 4261 GYTTTTCAGGAATCAAGACTCTTCACACCTGTGTGAATCTTAGACCCACGTTA 4320  
Oy 4322 CTGACCTGCTACCTCGAATTCACACCCAGATGGTGCAGATGCTGCTGATG 4381  
Db 4321 CTGACTGCTACCTTGAATTCACACCCAGATGGTGCAGATGCTGCTGATG 4380  
Oy 4382 GAGTGTCTGAGCTGAGGACAGGACCTGATG 4417  
Db 4381 GAGTGTCTGAGCTGAGGACAGGACCTGATG 4416

## RESULT 2

US-08-683-839B-2  
Sequence 2, Application US/08683839B  
Patent No. 5744326

## GENERAL INFORMATION:

APPLICANT: ILL, Charles . R. et al.  
TITLE OF INVENTION: Use of Viral Cis-Acting Post-Transcriptional  
TITLE OF INVENTION: Regulatory Sequences To Increase Expression of  
TITLE OF INVENTION: Intronsless Genes Containing Near-Consensus Splice Sites  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street, suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/683,839B  
FILING DATE: 11-MARCH-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:

## FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Remillard, Jane E.  
REGISTRATION NUMBER: 38,872  
REFERENCE/DOCKET NUMBER: TTI-138  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941

## INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
LENGTH: 9354 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 2965..7378  
US-08-683-839B-2

Query Match 99.8%; Score 4407.4; DB 1; Length 9354;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 4411; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 catgcaaatagagctctcaacgtctcttctggtcctttgagattcgtcttaagtc 60  
Db 2964 CATGGTTTATGAGCTCTCCACCTGCTTCTTCTGTGCTTTGGATTGCTTTAGTGC 3023  
Oy 61 caccagaataactactcctgggtgagtgagactgtcatggaactatgcaaatgagtc 120  
Db 3024 CACGAGAAGTACTACTCTGGGTGCTGAGTGAAGTGTCTATGGACTATATGCAAAAGTATCT 3083

Oy 121 cagtgagctgctgtgagcgaagatttccctctagagtgccaaatcttccaltcaa 180  
Db 3084 CGGTGACCTGCTGTGTGAGCCAGAGATTCTCTAGAGTGCAGAAATCTTTCCATCAA 3143  
Oy 181 cactcagtcgtgtacaaaagactcgtttgtagaattcaagttcaactttcaact 240  
Db 3144 CACCTCAGTGTGTACAAAAGACTGTTGTGAAATTCAGGTTTACCTTTTCAACAT 3203  
Oy 241 cgtcaagcgaagcgcacccctgagatggtctcgtagtgctccaccatcgaagctgagttta 300  
Db 3204 CGTAAAGCCAAAGGCGCACCTGAGATGGGTCTGTAGTGCCTACCATCAAGGCTGAGTTTA 3263  
Oy 301 tgatacagtggtcaatcaacttaagaacatggtcccatcctgtcaagttcaagtcgt 360  
Db 3264 TGATACAGTGTGTATACCTTAAAGAACATGGCTTCCCATCTGTGATAGTCTGAGTCTGT 3323  
Oy 361 tgggtatcctactcggaaagctctcgtgagagcttgaaatgatgatcaagaagcgttaag 420  
Db 3324 TGGTGTATCCTACTGGAACCTTCTGAGGAGCTGAAATGATGATCAGCCAGTAAAG 3383  
Oy 421 ggaagaagaatgataaagttctccctggttggaagccatataatgctgagagtcct 480  
Db 3384 GGAAGAAAGAAATGATATAAGTCTTCCCTGCTGGAAGCCATACATATGTCTGGAGGTCT 3443  
Oy 481 gaaagaagaatggtcccaatggtcccttgaccacatggtccttaactatcatcttctca 540  
Db 3444 GAAAGAAATGATGTCAATGAGCTGACCTGACCATGCTGCTTACCTCATATCTTCTCA 3503  
Oy 541 tgtgagacctgtgtaaaagacttgaaatcaggtccatgagagcctcactagtatgaga 600  
Db 3504 TGTGACCTGCTGTAAGAACTTGAAATTCAGGCTCATATGGAGCCCTCATGATGTAAGA 3563  
Oy 601 aggaagctgtgcccagaagaaagacacagacctgacaaatctataactcttctgt 660  
Db 3564 AGGAGTCTGCGCCAAAGAAAGAACACAGACCTTGCAAAATTAATATCTTTTCTCTG 3623  
Oy 661 atttgatgaagggaaagtgtggaactcagaacaaagaactccttcatgcaagtaggga 720  
Db 3624 ATTGATGAAGGGGAAAGTGTGCGACTCAGAAACAAAACTCTTATGATGAGATGGGA 3683  
Oy 721 tgcgcacatcgtcgtggtcgtggtcgttaaatgacacagatgattatgataaaggtc 780  
Db 3684 TGTGCAATCTGCTGGGCTGCGCTGAGCAATGACACAGTAAATGTTATGTAAACGGTCT 3743  
Oy 781 tctgcagcttgatgtgatgcccagaagaaatcagttatctgcatggtgattggaatgg 840  
Db 3744 TCTGCCAGTCTGTGATGTGAGTCCACAGGAATCATGTTATGGCAATGTGATGGATGG 3803  
Oy 841 caccactctgaagtgcactcaaatatctcctogaaggtcacacattctctgtgaggaaca 900  
Db 3804 CACCACCTCTGAAGTGCACATCAATATCTCTGAAGGTCAACACATTTCTGTGAGGAACA 3863  
Oy 901 tcgcgaagcgtccttggaatctgcgaataacttcttaactggtcgaacactcttgat 960  
Db 3864 TCGCCAGGCGCTTGGAAATCTGCCAATATCTTCTTATGCTCAAAACACTTGTGAT 3923  
Oy 961 ggaacctggaagtttctactgttttgcatatctcttccacaacaatgatgcatgga 1020  
Db 3924 GAGCCTTGAGACATGTTTCTACTGTTTGTATCTCTTCCACACATATGATGATGGA 3983  
Oy 1021 agcttatgtcaaatgtagcagcgtgtccagaggaaccccaactcgaatgaaaaataatga 1080  
Db 3984 AGCTTATGTCAAAATGACAGCTGTCCAGAGGAACCCCACTCATGAAATGAATATGA 4043  
Oy 1081 agaagcgaagactatgatgatcttactatggttctgaaatgagtgtgtcaagttta 1140  
Db 4044 AGAAGCGGAAGACTATGATGATGATCTTACTGATCTGAAATGATGATGCTGAGTTGA 4103  
Oy 1141 tgaatgaacacttctcctcttccatccaaatcgtcagttgccaagaagaagatcctaagc 1200  
Db 4104 TGATGACACACTCTCTCTTCTTATCCAAATTCGCTAGTTGCCAAGAGCATCTTAAAC 4163  
Oy 1201 ttgggtacattaatctgtcgtcgtgaagaagagactggactatgctccctttagtctcgc 1260

|||||  
Db 4164 TTGGATACATTTACATTTGCTGCTGAAGAGAGAGCTGGACTATGCTCCCTACTCTCGG 4223  
QY 1261 ccccgatgacagaagttataaagaatcaataattgaacaatgcccctcagcgagttgtag 1320  
Db 4224 CCCCAGTGAAGAGATTAATAAAGTCAATATTGAAACAATGGCCCTCAGCGGATTTGGTGT 4283  
QY 1321 gaagtaacaaaagtcggattatgcatcacagatgaaccccttaagaactgtaaac 1380  
Db 4284 GAAGTCAAAAAAGTCCGATTTATGCAATACAGATGAACCTTTAAGACTCTGTAAGC 4343  
QY 1381 taticgacatgaatcagaatccttggacccttacttctatgaggaaattggagacacac 1440  
Db 4344 TATTAGCATGAATCAGGAATCTTGGACCTTACTTTATGGGGAAGTTGGAACACACACT 4403  
QY 1441 gtgtatataattaaagaatcaagaacagacacataacaatcaccctcaaggaaatcac 1500  
Db 4404 GTTGATTATAATTAAACAAACCAAGACACATATACTACCTCCACGGAATCAC 4463  
QY 1501 tgatgcgcgtccctgttattcaaggagattaccaaagggtgaacacatttgaagatt 1560  
Db 4464 TGATGTCCTCTTCTTATTTCAAGGATTTACCAAAAGGTGAACATTTGAAGATTT 4523  
QY 1561 tccaattctgcagagaaatattcaaalataaattgacagtgactgtagaagatgggc 1620  
Db 4524 TCCAAATTCTGCCAGAGAAATATTCAAAATATAATGACAGTCACTGTGAAGATGGCC 4583  
QY 1621 aactaaataagatcctcgtgtgcctgacccgctatactactagtttcgttaataaggag 1680  
Db 4584 AACTAAATAGATTCCTCGGTGCTCGCTACCGCTATATCTAGTTTCTTAATATGAGAG 4643  
QY 1681 aagatcagcttcaagatcaattgagccctcctcactctgctacaagaactcgttaata 1740  
Db 4644 AGATTCAGCTTCAGGACGCTATTGGCCCTCTCTCATCTGCTAATAAAGATCTGTACATGA 4703  
QY 1741 aagaagaaacagataatgtcagacaagaagaaatgcatccgttcttctgtatattgata 1800  
Db 4704 AAGAGAAACCAAGATTAATGTACAGACAGAGAAATGTCATCCGTGTTTCTGATTTATGA 4763  
QY 1801 gaacgaagctgttacctacacagaaataacaagccttctcccaatccagctgagat 1860  
Db 4764 GAACCAACCTGTGTACTCTACAGAGAAATATACAAACCTTCTCCCAATCCAGCTGAGAT 4823  
QY 1861 gaagcttggagatccagagattccaaagcctcaaatcatgacaagataaaggcattgt 1920  
Db 4824 GCAAGCTTGAGATCCAGAGTTCCAAAGCTTCACATCATATGACACAGATCAATAGGCTATGT 4883  
QY 1921 ttctgaatgttgcagttgttcagttgtttgtcatgagtgagcatctggtlacaatcga 1980  
Db 4884 TTTTGATAGTTGCAAGTGTCAAGTTGTTTGATGATGAGTGGCATACTGGTACATTTCTAAG 4943  
QY 1981 catitggacacagactgactccttctcgtctctctctctgtatataccttcaaacaa 2040  
Db 4944 CATTTGGACACAGACTGACTTCTCTCTCTCTCTCTCTCTGATATACCTTCCAAACACAA 5003  
QY 2041 aatgtgctatgaaac 2100  
Db 5004 AATGGCTATGAAAGACACACTACACCTATTTCCATTTCTAGAGAAACTGCTCTTCAATGTC 5063  
QY 2101 gatggaaaaacacagttctatgattcttgggtgtccacaactcagacttgcgaaacagag 2160  
Db 5064 GATGAAAAACCCAGGTCATGATTTCTGGGGTGGCCCAACTCAGACTTTCGGAACAGAG 5123  
QY 2161 catgacgccttcaatcgaaggttctcagttgtgacaagaacactggtgattatacgaaga 2220  
Db 5124 CATGACCGCTTACTGAAAGGTTTCTAGTTGTGACAAAGAACACTGGGATTTATTACAGAA 5183  
QY 2221 cagttatgaagatatttgcagacttctgtgagtaaacaatgcatatgaaacaaagag 2280  
Db 5184 CAGTTATGAAGATATTTCAGCATACTTGTGTGATTAATAAATGCCATTTGAACCAAGAAAG 5243  
QY 2281 ctctctcccaaatctcaagacacccctagcactagcaaaaagcaatlaatgcaacccacac 2340  
|||||

Db 5244 CTTCCTCCAGAAATTTCAAGACACACCTTAGCATAGCAAAAAAGCAATTTAATGCCACCCACC 5303  
QY 2341 agtcttgaagacgcacatcaacgggaataactcgttactactctcagtcagatcaagaga 2400  
Db 5304 AGCTTGAAACGCCATCAACGGGAAATTAACCTGCTACTACTCTTCACGTCAAGATCAAGAGA 5363  
QY 2401 aattgactatgatatcaatcatcagtttgaagaagaagaatttgaacttataga 2460  
Db 5364 AATTGACTATGATGATATCCATATACATTGAATGAAGAAAGAAATTTTGACATTTATGA 5423  
QY 2461 tgaaga tgaanaalcagagcccccagccttccaagaanaacacagacatttata tgc 2520  
Db 5424 TGAGATGAAAAATCAGAGCCCCCGCAGCTTTCAAAAGAAAACGACACTATTTATTGC 5483  
QY 2521 tgcagtggagaggctctggattatagggatgagtaatcccccaatgcttcaagaacag 2580  
Db 5484 TGCAGTGGAGAGCTCTGGATTTATGGGATGATGATACCTCCCAATGTTCTAAGAAAACAG 5543  
QY 2581 ggtcagagtgagtgagtgctccctcagttcaagaagaatggtttccaggaatttactgatg 2640  
Db 5544 GGCTCAGATGGCAGTGTCCCTCAGTTCAAGAAAGTGTGTTTCCAGGAATTTACTGATGG 5603  
QY 2641 ctcccttactcagcccttataaccgttggagaactaaatgaacatttggactcctgggccc 2700  
Db 5604 CTCCCTTACTCAGCCCTTATACGTGGAGAACTAATGAACATTTGGAGACTCCTGGGGCC 5663  
QY 2701 atataagaagcagaagtttgaagaataatcatatgtaacttcaagaatcaggcctctcg 2760  
Db 5664 ATATATAAAGCAGAAAGTTGAAGATATATCATAGTAATCTTCAGAAATCAGGCTCTCG 5723  
QY 2761 tccctatctctcttactcagccttattcttctatggagaagatcagaaggcaagaagaca 2820  
Db 5724 TCCCTTATCTCTTATTTATTTAGCCTTATTTCTTATGAGAGATCAAGCAAGCAGAGACAG 5783  
QY 2821 acctagaaaaaacttltgcaagccctlaatgaacccaaacttacttltggaagatgcaaca 2880  
Db 5784 ACTAGAAAAAATTTGTCAAGCCTAATGAACCAAACTTACTTTGGAAATGCAACA 5843  
QY 2881 tcaatggaacccacataaagaatgttgaatgcaagcctgggcttattctctgagt 2940  
Db 5844 TCATATGGACCCACCTAAAGATGAGTTGATGCTCAAGCAAGCCTGGCTTATTTCTCATGTG 5903  
QY 2941 tgaactggaaaaaagatgttgcactcagccttgaattggaacccctcgtgtctgcacactaa 3000  
Db 5904 TGACCTGGAANAAGATGTGCACCTCAGGCTGATTTGGAACCCCTTCTGTGCTGCACACTAA 5963  
QY 3001 cacactgaaacccctgtcatcagtggaagaatgacaagtacaggaatttgcctcgtgttccac 3060  
Db 5964 CACACTGAACCCCTGCTCATGGGAGACAAGTGAACAGTAAGAAATTTGCTGTGTTTTCAC 6023  
QY 3061 catcttltgattggaccacaagaactcgtgaacttcaatgaataatggaagaactgcagggc 3120  
Db 6024 CATCTTTGATGAGACCAAAAGCTGTGACTGCACTGAAATATGGAAGAAACGACGGGAC 6083  
QY 3121 tccctgcaatccagatggaagaatcccaacttlaagaagaatttactcgttccatcaat 3180  
Db 6084 TCCCTGCAATATCCAAATGGAAGATCCCACTTTTAAAGAAATTAACGCTTCATCAAT 6143  
QY 3181 caatggtacataatgatalacactactcgtgttagtaatggtcagatcaaaagattcg 3240  
Db 6144 CAATGCTCATATATGATGATACACTACTGCTGTAGTAATGCGTCAGAGATCAAAAGGATTCG 6203  
QY 3241 atgtgatactgtccagacatggtgagcaatgaanaacatccattctattcatltcaagtga 3300  
Db 6204 ATGTATCTGTCTCAGATGGCAGCAATGAANAACATCAATTTCTATTTCAATTCAGTGA 6263  
QY 3301 tgtgttcaactgttgaagaaaaaagagagatataaattggagactgtacaatcctatcag 3360  
Db 6264 TGTGTTCACTGTACGAAAAAAGAGAGATTAATAATGCACTGTACATCTCATACAG 6323  
QY 3361 tgttlttggacagtgaaatgttacaatccaaagcttgaatttggcgggttgaatgact 3420  
Db 6324 TGTTTTGGACAGTGAAGATGTTACATCCAAAGCTGGAATTTGGCGGTGGAATGCT 6383  
|||||



OY 3421 tatltggcgcagcaltacatcgtctggagatgagcacactttcttggltgltacagcaatagtg 3480  
|||||  
DB 6384 TATTGGCAGCATCTACATCTCGAGATGAGCACACTTTTCTGTGTATACACAAATAGTG 6443  
OY 3481 tcagactcccctgggaatggtctctggacacatlaagatatttcagattacagttcagg 3540  
6444 TCAGACTCCCTGGGAATGGCTTCTGGACACATTAGAGATTTCAGATTACAGCTTGAAG 6503  
OY 3541 acaataatgagcagtgaggcccaagcctggccagactcatattccggaatcaatcagtc 3600  
6504 ACATATATGACAGATGGGCCCCCAAGCTGGCCAGACTTATTATTCCGATATCAATATGC 6563  
OY 3601 ctggagcaccnaagagaccccttcttcttggatcaagtgatctgtctggcaccatgatat 3660  
6564 CTGGAGCACCNAAGAGACCCCTTTCTTGTGATCAAGTGTGATCTGTGGACCAATGATTAT 6623  
OY 3661 tcaagcgtcaagaaccgaaggtgcggctcagaagttctcagactctcaactctcagtc 3720  
6624 TCACGGCATCAAGACCCAGGGTCCCGTCAGAAATTCTCCAGCTTCAATCTCAGTT 6683  
OY 3721 tatcatcatatagctctgtatgagaaagtgagcagactatcgaagaaatctcactgg 3780  
6684 TATCATCATATAGTCTTGTATGGAAAGAGTGGCAGACTTATGAGAAATTTCCACTGG 6743  
OY 3781 aacctaatgtctcttcttggcaatgtgatalcatctggataaacaacataattttaa 3840  
6744 AACCTTAATGTCTCTTGTGCAATGTGATCATCTGGGATAAACAATATTATTAA 6803  
OY 3841 cccctcaattatctgtcgtacataatcgttctgcaaccaactcatatagcattgcagac 3900  
6804 CCTTCAATATTGTGCGATACATCCGTTTGACCCCAACTATTATGCAATTCACACAC 6863  
OY 3901 tcttcgcatgaggttgatgggctgtgatttaaatagttgcagcaatgcatatggagag 3960  
6864 TCTTCCATGAGATGTATGGGCTGTGTATTAATAGTTGCAAGCTTGGCAATGGAAATGGA 6923  
OY 3961 ggttaagaacataatcagatgacagatlaactgtcttcaacttcaacataatgtttgc 4020  
6924 GAGTAAAGCAATATCATGATSCACAGATTACTGCTCATCTACTTACCAATATGTTTGC 6983  
OY 4021 cactgtgtctcttcaaaaagctcgtacttcaactcccaaggagagatagtcctggagac 4080  
6984 CACTGTGTCTCTTCAAAAAGCTCGACTTCACTTCAAGGAGAGTATATGCTTGAGACC 7043  
OY 4081 tcaagtgaaataatccaaaagatggtcgtcgaagtgtgacttccagaagaacatgaagtcac 4140  
7044 TCAGGTGAATTAATCCAAAAGATGGCTGCAAGTGGACTTCCAGAAAGCAATGAAAGTCAC 7103  
OY 4141 aggaagtaactactcaaggagtaaaatctctgtcttcaacagatgatagtgaagagttcct 4200  
7104 AGGAGTAATCTACTCAGGAGTAAATCTCTGCTTACCCAGCATGATGTGAAGGATTCT 7163  
OY 4201 catctcagaagcagaatagggcatcagtagtactcttcttcaagaatgcaagaataaa 4260  
7164 CATCTCCACAGTCAAGATGGCCATGAGTGAAGTCTCTTTTTCAGAAATGGCAAAAGTAAA 7223  
OY 4261 ggttttccagggaacaaagaccccttcaacactgtgtgaactctctagaaccacgctc 4320  
7224 GGTTTTTCAGGGAAATCAAGACTCTTTCACACTGTGTGAATCTCTAGACCCACCGTT 7283  
OY 4321 actgactcgtactactcgaattcaaccocagagltgggtgacacagatgcccctgaagat 4380  
7284 ACTGACTCGTACTCTGCAATTCACCCCGAGATGGGTGACACAGATTCCTCCGAGAGAT 7343  
OY 4381 ggaagttctcgggtcgcagagcaagacactctactga 4417  
7344 GGAGTTCTGGGCTGCGAGGACAGGACCTCTACTGA 7380

RESULT 3  
US-08-717-294-41  
; Sequence 41, Application US/08717294

Patent No. 6114148  
: GENERAL INFORMATION:  
: APPLICANT: SEED, BRIAN  
: APPLICANT: HAAS, JURGEN  
: TITLE OF INVENTION: HIGH LEVEL EXPRESSION OF  
: NUMBER OF SEQUENCES: 110  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Clark & Elding LLP  
: STREET: 176 Federal Street  
: City: Boston  
: STATE: MA  
: COUNTRY: USA  
: ZIP: 02110  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Diskette  
: COMPUTER: IBM compatible  
: OPERATING SYSTEM: DOS  
: SOFTWARE: FASTSEQ for Windows Version 2.0  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/717,294  
: FILING DATE: 20-SEP-1996  
: CLASSIFICATION: 435  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER:  
: FILING DATE:  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Elding, Karen L.  
: REGISTRATION NUMBER: 35,238  
: REFERENCE/DOCKET NUMBER: 00786/345001  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 617-428-0200  
: TELEFAX: 617-428-7045  
: TELEX:  
: INFORMATION FOR SEQ ID NO: 41:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 4670 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: double  
: TOPOLOGY: linear  
: MOLECULE TYPE: cDNA  
US-08-717-294-41  
Query Match 98.7%; Score 4357.4; DB 3; Length 4670;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4358; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 59 gccacccaagaatactactcctgggtgcagtggaactgtcatggaactatatgcaagtgat 118  
DB 84 GCCACCCAGAAGATACATCTCGGCTGCAGTGGAACTGTCATGGGACATATGCAAACTGAT 143  
OY 119 ctgggtgagctgctgtggcgcaagaattcctcctcagaagtgccaaatcttttcaattc 178  
DB 144 CTGGGTGAGCTGCTGTGGAGCGCAAGATTTCCTCTAGAGTGCAAAATCTTTTCATTC 203  
OY 179 aacacctcagtcgttcaaaaagactcgttgtgaaattcaaggttcaacttttcaac 238  
DB 204 AACACCTCAGATCGTGTCAAAAAGACTCTGTTGTGAATTCACGGATCACCTTTTCAAC 263  
OY 239 atcgctaagccaaggccacccctggatggtgtcgtctaggtgtccatccaatcgaagctgagtt 298  
DB 264 ATCGCTAAGCCAAAGGCCACCCTGGATGGGTCTGTAGTCTTACCATTCACAGGCTGAGTT 323  
OY 299 tatgatacagtgatcttaacttaagaacatggtcttccatctctgaagttcagttcagtt 358  
DB 324 TATGATACAGTGTGTCATTTAAGAAACATGGCTTCCATCTGTGCACTTCAATGCT 383  
OY 359 gttggttaccctacgtggaagcttctgagggagctgaaatagatgataagacaaatcga 418  
DB 384 GTTGCTTATCTACTAGGAAGCTTCTGAGAGAGCTTGATATATATATATGACACATGCA 443  
OY 419 agggagaagaagatgataaagttcttccctcgtgtggaagccatacatatgtctggcagtc 478

|||||  
Db 444 AGGAGAGAAAGATGATATAAGTTCCTCGTGGAGGACCATACATATGCTGGCAGGTC 503  
QY 479 cttaaaagaatggtccaatgagcctcagaccctgctcttaactaatactcttc 538  
Db 504 CTGAAAGAGATGGTCCAAAGGCGCTCGACCCACTGTCCTTACCTATATCTTTCT 563  
QY 539 catgtgacctggttaaaagacttgaatcagacctcaatgagccctactagataga 598  
Db 564 CATGTGACCTGGTAAAGATTGAATTCAGGCTCTATTGGAGCCCTACTAGTATGAGA 623  
QY 599 gaaggagctctggccaagaaagaaacagaccttgacaaatttactacttctgt 658  
Db 624 GAAGGGAGTCTGGCCAAAGGAAAGACACAGACCTTGACAAATTTATCTATTTTGTCT 683  
QY 659 gttattgataagaaagatttgacctcagaaacaaagaccctctgtatgcagatagg 718  
Db 684 GTATTGTGATGAAGGAAAGTTTGGCACTCGAAGAAACAAAGAACTTTGATGACGATAGG 743  
QY 719 gatgtgcatctgtctgggacctgacctaaatgacacaatgcatgttattgtaaacag 778  
Db 744 GATGCTGCATCTGCTGGGCGCTGGCCTAAATGCAACAGTCAATGGTTATGTAAACAGG 803  
QY 779 tctctgccaagttctgattggaatgccaaggaatcagttctatgcatgttattggaatg 838  
Db 804 TCTGTGCAGAGTCTGATTTGATGATGCCACAGGAATCAGTCTATTGGCATTTGGATG 863  
QY 839 ggcacacactcctgaaatgcatcaataatcctcctgaaggtcacaacattctgtggaagac 898  
Db 864 GGCACACACTCTAAATGCACTCAATAATTCCTGGAAGGTGCACAAATTTCTTGGAGGAAC 923  
QY 899 cactgcgaagcgtctcttggaatctcgccaataacttctctactgctcacaacaccttg 958  
Db 924 CATCGCAGGCGCTCTTGGAATCTGCCAATAATCTTCCCTTACTGCTCAAAACCTTTC 963  
QY 959 atggaccttggaagttctactgtttgtgcatactcttcccaacaatgattggaatg 1018  
Db 984 ATGACCTTGGAGAGTTCCTACTGTTTGTGCATATCTCTCCACCAACATGATGCGCATG 1043  
QY 1019 gaagcttatgtcaaaatgaagcagctgtccagagaaacccaactgaatgaaataat 1078  
Db 1044 GAAGCTTATGTCAAGATGACAGACTGTCCAGAGAAACCCCACTACGAATGAAGAAATAAT 1103  
QY 1079 gaagaagcgaagactatgataatgatacttactgacttctgaaatggaatggtcaggtct 1138  
Db 1104 GAAGAAGCGGAACATATGATGATGATCTTACTGATTTGGAATGATGATGGTGCAGGTTT 1163  
QY 1139 gatgatgacaactctcctcttactltaacaatcgctcagttgccaagaagacatccataa 1198  
Db 1164 GATGATGACAACTCTCTCTCTTATTCCAAAATTCGCTCAGTTGCCAAGAAAGCATCTTAA 1223  
QY 1199 acttggtatcatcaactgtctgtctgaagaggaagactgggactatgctcctctagtcttc 1258  
Db 1224 ACTTGGGTACATTACATTTCTCTGTAAGAGGAGGACTGGGACTATATCTCCCTTAGTCTTC 1283  
QY 1259 gcccccgaagacgaagttataaaatgcaatattgaaacaatggccctcagcgatgtgt 1318  
Db 1284 GCCCCGATGACGAAGTTRTAAAGTCAATATTGTAACAATATGGCCCTCAGCGGATTTGGT 1343  
QY 1319 aggaagttacaaaaaagttccgattatgcatcacacagatgaaaccttlaagactctgaa 1378  
Db 1344 AGCAATGACAAAAGTCCGATTTATGCAATACAGATGAACCTTTAAAGACTCTGTAA 1403  
QY 1379 gctatcagcatgaatcaagaaatcttggagaccttactcttattggggaagttggaagaca 1438  
Db 1404 GCTATTTCAGATGATCAGGAATCTTTGGAGCCTTTTACTTTATGGGAGTTGGAGACACA 1463  
QY 1439 ctgttatattatlaagaatcaagaagaagacatatlaacatccacctcaagaaac 1498  
Db 1464 CTGTTGATTTATTTTAAGATTCAGCAAGCAGACCATATTAACCTTCAACCTCAGGAATC 1523  
QY 1499 actgatgtccgtctctgttattccaagagattaccaaaagttgtaaacattgaaagat 1558  
|||||

Db 1524 ACTGATGTCCCTCTTTGTATTCAGAGATATACCAAAAGTGTAAACATTTGAAGAT 1583  
QY 1559 ttccaattctgccaagagaataatccaataataatgacagttgactgtaagaagttgg 1618  
Db 1584 TTTCCAAATTTCTGCAGAGAAATATTCAAATATTAATGACAGTGCATGAGAAATGGG 1643  
QY 1619 ccaactaaatcagatcctctgtctgacccgcatatactctactgttctgttaataatggag 1678  
Db 1644 CCAACTAAATCAATCTCTCGTGCCTGACCCGCTATTACTCTAGTTTCGTAATATGGAG 1703  
QY 1679 agagatcctagcttcaagactcaatlggacctctctcatcctgtctacaaagaatctgatat 1738  
Db 1704 AGAGATCTAGCTTCAAGACTCATTTGGCCCTCTCTCATCTGTACAAAGAAATCTGTAGAT 1763  
QY 1739 caaagggaaaccaaataatgtcagacaagagaatgcatcctgttctctgtatttgtat 1798  
Db 1764 CAAGAGGAAACCGATTAATGTGACAGAGAAAGAAATGTCATCTGTATTGTGAT 1823  
QY 1799 gagaaccgaagctggttaacctcacagagaataatacaagcttctcccaatccagcttga 1858  
Db 1824 GAGAACCGAAGCTGTGACTTCCACAGAGAAATATCAACGCTTTTCCCCANTCCAGCTGGA 1883  
QY 1859 gtgcaacttgaggaatccagagttccaaagctccaaacatcatgacagacatcaatggctat 1918  
Db 1884 GTGACACTTGAGGATCCAGAGTTCCAAAGCCTCCAAACATCATGCAAGCATGATGCTAT 1943  
QY 1919 gtttttgatagtttcagttgcatgttctgttgcataaggttgacatacttca 1978  
Db 1944 GTTTTGTATAGTTTGAGTTGATGATGTTGTTGTCATAGAGTGGCAATCGTCAATTCAT 2003  
QY 1979 agcatlggaagacagactactcctctctctctctctctctctctctctctctctctcaaac 2038  
Db 2004 AGCATTTGAGACACAGACTGACTCTCTTCTGTCTTCTTCTGTGATATACCTTCAACAC 2063  
QY 2039 aaaaatgctatgaagaacaactacacctatcccatctcagagagaacatgctctcatg 2098  
Db 2064 AAAATGCTATGAAAGACACACTCACCTATTCCTCTTCTCAGAGAAACCTGCTTCATG 2123  
QY 2099 tgaatgaaacccaagttcatgtatgtctgtgggtgccaacactcaacttctggaacaga 2158  
Db 2124 TCGATGAAACCCAGGTCATGATGATGCTGGGTGCGACAACTCAACATTCGGAACAGA 2183  
QY 2159 ggcataccgctctactgaaggttctctctgtgcaagaacaactggtgattatgaag 2218  
Db 2184 GCGATACCGCGCTTACTGAAAGTTTCTAGTTGTGCAAGAACCTGTGTATTTACGAG 2243  
QY 2219 gacagttatgaagataltccaagactctgtgtgagtaaaaaaacaatgccaatgaaacaga 2278  
Db 2244 GACAGTTATGAAGATATTTACAGCATCTTGTGAGTAAAAACAATGCCATTTGAACCAAGA 2303  
QY 2279 agcttctcccaagaatccaagaacacctagcactaggcaaaagaaatttaatgcaccca 2338  
Db 2304 AGCTTCTCCAGAAATTCAGAGACACCTTACACATGAGCAAAAGCAATTTAATGCCACCA 2363  
QY 2339 ccagctctgaaacgcatcaacgggaataaactcgtactactcttcaagtcagaatgaag 2398  
Db 2364 CCACTTTTAAACGCCATTAACGGGAATTAACCTGACTACTCTTATAGTCAGATTCAGAG 2423  
QY 2399 gaaattgactatgataccaatalcagttgaaatgaaagaaagaaatttgacattat 2458  
Db 2424 GAAATGACTATGATATACATATCATGTTGAATGAAGAAAGAAAGATTTTGACATTTAT 2483  
QY 2459 gatgaggaatgaataatcagagcccccgcaacttcaaaaagaaacagacaactatttatt 2518  
Db 2484 GATGAGGATGAATAATAGAGCCCCCGACCTTTCAAAAACAAAGAACATATTTTATTTAT 2543  
QY 2519 gctgcagttgagaagctctcgggattatgggaatgaaggtgtcccaaatcttcaagaagac 2578  
Db 2544 GCTGCGATGAGAGGCTCTGGGATTTATGGGATGATAGTCTCCCCACATATTTCTAAGAAAC 2603  
QY 2579 agggctcagaatgagcagttctcctcagttccaagaaatgttcttccaagaaattactgat 2638  
Db 2604 AGGGCTCAGAGTGAGTGTCTCCTCAGTTCAAGAAAGTTGTTTCCAGGAATTTACTGAT 2663  
|||||

QY 2639 ggcctcttactcagcccttatacagctggagaactaaatgacatttggagctctcggg 2698  
|||||  
Db 2664 ggcctcctttactcagcccttatacagctggagaactaaatgacatttggagctctcggg 2723  
QY 2699 ccaataataagagcagaagtgtgaagataatacaatgtaacttccaagaatcaggctctc 2758  
|||||  
Db 2724 ccaataataagagcagaagtgtgaagataatacaatgtaacttccaagaatcaggctctc 2783  
QY 2759 cgtccctactctctctacttctagccttattctctatagagaagatacagaagcagaagaca 2818  
|||||  
Db 2784 cgtccctactctctctacttctagccttattctctatagagaagatacagaagcagaagaca 2843  
QY 2819 gaacctgaagaaaaacttctgaagccttaataagaaacaaacttacttctggagaagtgcga 2878  
|||||  
Db 2844 gaacctgaagaaaaacttctgaagccttaataagaaacaaacttacttctggagaagtgcga 2903  
QY 2879 caataatggaacccactaaagaatgtaactgtaactgcaaaagccctggcttattctctgat 2938  
|||||  
Db 2904 caataatggaacccactaaagaatgtaactgtaactgcaaaagccctggcttattctctgat 2963  
QY 2939 gtggaacttgaaaaagaatgtgcaactcagccttgaatggaccctcttggtctgcacact 2998  
|||||  
Db 2964 gtggaacttgaaaaagaatgtgcaactcagccttgaatggaccctcttggtctgcacact 3023  
QY 2999 aacacactgaacccctgtctatgggagaacagtgaacagtaacagaatttgcctctgtctc 3058  
|||||  
Db 3024 aacacactgaacccctgtctatgggagaacagtgaacagtaacagaatttgcctctgtctc 3083  
QY 3059 aacacacttgtaagaaacaaagctgtgtaactcactgaagaaataatgagaagaactgcagg 3118  
|||||  
Db 3084 aacacacttgtaagaaacaaagctgtgtaactcactgaagaaataatgagaagaactgcagg 3143  
QY 3119 gccctctgaacataatccagatcgaagatcccaactttaaagaagatttcgtctccatgca 3178  
|||||  
Db 3144 gccctctgaacataatccagatcgaagatcccaactttaaagaagatttcgtctccatgca 3203  
QY 3179 atcaatggtacataatgataatgatacactggtctgaatgtaactgcaagatcaaaagatt 3228  
|||||  
Db 3204 atcaatggtacataatgataatgatacactggtctgaatgtaactgcaagatcaaaagatt 3263  
QY 3239 cgaatgtaactgctcagcagcagcagcagaatgaagaacatccatctatctatcttaagga 3288  
|||||  
Db 3264 cgaatgtaactgctcagcagcagcagcagaatgaagaacatccatctatcttaagga 3323  
QY 3299 catgtctcactgtacgaagaaagagagataataatggaactgtaacatctctatcca 3358  
|||||  
Db 3324 catgtctcactgtacgaagaaagagagataataatggaactgtaacatctctatcca 3383  
QY 3359 ggtgtcttctgagacagbgaatgtaacatccaaagctggaaattggcgagggaatgc 3418  
|||||  
Db 3384 ggtgtcttctgagacagbgaatgtaacatccaaagctggaaattggcgagggaatgc 3443  
QY 3419 ctatattggcagagatctactactgctggagatgagacacatttctcgtgtgaagaataag 3478  
|||||  
Db 3444 ctatattggcagagatctactactgctggagatgagacacatttctcgtgtgaagaataag 3503  
QY 3479 tctcagaactccctgggaatgagctctcgaacacatagagaatttccagatacagcttca 3538  
|||||  
Db 3504 tctcagaactccctgggaatgagctctcgaacacatagagaatttccagatacagcttca 3553  
QY 3539 ggaacaaatagagacagtgggcccacaaagctggcgaacttcaatctccagatacacaat 3588  
|||||  
Db 3564 ggaacaaatagagacagtgggcccacaaagctggcgaacttcaatctccagatacacaat 3623  
QY 3599 ggcctgagagaacagaagcccttctctgtaagaagtgatcgtgtggaacaaatatt 3658  
|||||  
Db 3624 ggcctgagagaacagaagcccttctctgtaagaagtgatcgtgtggaacaaatatt 3683  
QY 3659 attcaagagatacaagaacccagagtgcccgctcaagaagtctcccaagcctcactacatccag 3718  
|||||  
Db 3684 attcaagagatacaagaacccagagtgcccgctcaagaagtctcccaagcctcactacatccag 3743

QY 3719 ttatcatalcagatagatcttctgagtggaagaagtggcagaacttatacgaagaaattccact 3778  
|||||  
Db 3744 ttatcatalcagatagatcttctgagtggaagaagtggcagaacttatacgaagaaattccact 3803  
QY 3779 ggaaccttaatggtctcttcttggcaatggtgatctcctgggataaacaacatattttt 3838  
|||||  
Db 3804 ggaaccttaatggtctcttcttggcaatggtgatctcctgggataaacaacatattttt 3863  
QY 3839 aacctccaattatctctcgaatacaccgtttgacccaactcaataacatctgcagc 3898  
|||||  
Db 3864 aacctccaattatctctcgaatacaccgtttgacccaactcaataacatctgcagc 3923  
QY 3899 aacctcgaatgagctgagtgagctgtaatttaataatgtgcagatccatctgggaatg 3958  
|||||  
Db 3924 aacctcgaatgagctgagtgagctgtaatttaataatgtgcagatccatctgggaatg 3983  
QY 3959 gagaagtaagaataatacagatgacaaatattacatgcttacccttaccatattgctt 4018  
|||||  
Db 3984 gagaagtaagaataatacagatgacaaatattacatgcttacccttaccatattgctt 4043  
QY 4019 gccacctgctctctcgaagaagctcgaactcgaactcccaagggagagtaatgctcgaga 4078  
|||||  
Db 4044 gccacctgctctctcgaagaagctcgaactcgaactcccaagggagagtaatgctcgaga 4103  
QY 4079 cctcaggtgaataatccaaagaagtgctgcgaagtggacttccaagaagacaatgaaagtc 4138  
|||||  
Db 4104 cctcaggtgaataatccaaagaagtgctgcgaagtggacttccaagaagacaatgaaagtc 4163  
QY 4139 acaagagtaactactcagggagagtaaaatctctctaccagcagatgtaatggaaggctc 4198  
|||||  
Db 4164 acaagagtaactactcagggagagtaaaatctctctaccagcagatgtaatggaaggctc 4223  
QY 4199 ctcatctcagaacatcagaatgagcagatggaactcctcttctcagaatgcaaaagta 4258  
|||||  
Db 4224 ctcatctcagaacatcagaatgagcagatggaactcctcttctcagaatgcaaaagta 4283  
QY 4259 aaggttctcagggaaatcaagactcctcacaacgtggtgtaactctagaaccacg 4318  
|||||  
Db 4284 aaggttctcagggaaatcaagactcctcacaacgtggtgtaactctagaaccacg 4343  
QY 4319 ttactgactcgaactcctgaatcaccacccaagtggtgtaacagaatgtgcctcgaag 4378  
|||||  
Db 4344 ttactgactcgaactcctgaatcaccacccaagtggtgtaacagaatgtgcctcgaag 4403  
QY 4379 atggaagttctggtctgcagagcagaagactctactga 4417  
|||||  
Db 4404 atggaagttctggtctgcagagcagaagactctactga 4442

RESULT 4  
US-09-470-618-14  
: Sequence 14, Application us/09470618  
: Patent No. 6200560  
: GENERAL INFORMATION:  
: APPLICANT: Coult, Linda B.  
: APPLICANT: Coult, Peter C.  
: TITLE OF INVENTION: Adeno-Associated Vectors for Expression of Factor VIII  
: FILE REFERENCE: Avigen-04082  
: CURRENT APPLICATION NUMBER: US/09/470,618  
: CURRENT FILING DATE: 1999-12-22  
: EARLIER APPLICATION NUMBER: 09/364,862  
: EARLIER FILING DATE: 1999-07-30  
: EARLIER APPLICATION NUMBER: 60/125,974  
: EARLIER FILING DATE: 1999-03-24  
: EARLIER APPLICATION NUMBER: 60/104,994  
: NUMBER OF SEQ ID NOS: 15  
: SOFTWARE: Patentlin Ver. 2.0  
: SEQ ID NO 14  
: LENGTH: 4999  
: TYPE: DNA  
: ORGANISM: Artificial Sequence

|                       |              |               |               |                   |
|-----------------------|--------------|---------------|---------------|-------------------|
| Query Match           | 97.8%;       | Score 4319.8; | DB 4;         | Length 4999;      |
| Best Local Similarity | 99.0%;       | Pred. No. 0;  |               |                   |
| Matches 4373;         | Conservative | 0;            | Mismatches 2; | Indels 42; Gaps 1 |

|    |      |  |      |
|----|------|--|------|
| OY | 961  | ggaaccttggacaagtttctactgtttgttgcataatctcttcccaacaatagtatgcatgta      | 1020 |
| Db | 1375 | ggaccttggacagagtttctactgtttgttcataatctcttcccaacaatagtatgcatgta       | 1434 |
| OY | 1021 | agcctatgtcaaaagttagacagctgtctccagaaggaaaccccaactacagaaatgaaaataatga  | 1080 |
| Db | 1435 | agcctatgtccaagttagaacagcgtgtccagaaggaaaccccaactacagaaatgaaaataatga   | 1494 |
| OY | 1081 | agaagcggaagaaactatgatgatcatcttaagcttcttgaaatgtagtgtgtcgaagtttga      | 1140 |
| Db | 1495 | agaagcggaagaaactatgatgatcatcttctgaagcttcttgaaatgtagtgtgtcgaagtttga   | 1554 |
| OY | 1141 | tgatgtacaactctctctcctttatccaagtctgcagttgtccaaagaaacatccataaac        | 1200 |
| Db | 1555 | tgatgtacaactctctctcctttatccaagtctgcagttgtccaaagaaacatccataaac        | 1614 |
| OY | 1201 | ttaggtatcatctaacttgtctgtctgaagaaggaaacttggagctatgtctccttagtctctgc    | 1260 |
| Db | 1615 | ttaggtatcatctaacttgtctgtctgaagaaggaaacttggagctatgtctccttagtctctgc    | 1674 |
| OY | 1261 | ccccgaatgacagaagatgtataaagaatcaaatatttggacaatctggccctcagcgaattgtgag  | 1320 |
| Db | 1675 | ccccgaatgacagaagatgtataaagaatcaaatatttggacaatctggccctcagcgaattgtgag  | 1734 |
| OY | 1321 | gaagttacaaaaaagtcocgattttatgtgcatacacagaatgaacctttaaagactgtgaaagc    | 1380 |
| Db | 1735 | gaagttacaaaaaagtcocgattttatgtgcatacacagaatgaacctttaaagactgtgaaagc    | 1794 |
| OY | 1381 | tattcagcatgaaatccaggaatctcttggaaacttacttttgggggaagtgttggagaaacact    | 1440 |
| Db | 1795 | tattcagcatgaaatccaggaatctcttggaaacttacttttgggggaagtgttggagaaacact    | 1854 |
| OY | 1441 | gttgattatatttaagaaatccaagcaagcagacataataacatctacccctcaagcagatcac     | 1500 |
| Db | 1855 | gttgattatatttaagaaatccaagcaagcagacataataacatctacccctcaagcagatcac     | 1914 |
| OY | 1501 | tgatgtccgtctccttctgtatactcaagsgaataccaanaaggtgtanaaacatttgaaggaatt   | 1560 |
| Db | 1915 | tgatgtccgtctccttctgtatactcaagsgaataccaanaaggtgtanaaacatttgaaggaatt   | 1974 |
| OY | 1561 | tccaattctgcagggagaanaattatccaataataatgtgaagtgactgttagaagaatgggccc    | 1620 |
| Db | 1975 | tccaattctgcagggagaanaattatccaataataatgtgaagtgactgttagaagaatgggccc    | 2034 |
| OY | 1621 | aacccaatcagatctccgtggtcctggcccgcaatgaactcttagtttcogtttaatatgtgagag   | 1680 |
| Db | 2035 | aacccaatcagatctccgtggtcctggcccgcaatgaactcttagtttcogtttaatatgtgagag   | 2094 |
| OY | 1681 | agatcctagcttcagagactcaatttggccctctcctcatctgtctacaagaagaatcgttagatca  | 1740 |
| Db | 2095 | agatcctagcttcagagactcaatttggccctctcctcatctgtctacaagaagaatcgttagatca  | 2154 |
| OY | 1741 | aagggggaacccagataatgtctcaagaacaagggaatgtcatcccggttttctgtatatttgaga   | 1800 |
| Db | 2155 | aagggggaacccagataatgtctcaagaacaagggaatgtcatcccggttttctgtatatttgaga   | 2214 |
| OY | 1801 | gaaccggaagctgtggaaccttccacagagaataatacaagcgttctccccaatccagctcgaggt   | 1860 |
| Db | 2215 | gaaccggaagctgtggaaccttccacagagaataatacaagcgttctccccaatccagctcgaggt   | 2274 |
| OY | 1861 | ggcagcttgaagatccagagatttccaagcctctccaacatcatatgcaagcatccaatbtgctatgt | 1920 |
| Db | 2275 | ggcagcttgaagatccagagatttccaagcctctccaacatcatatgcaagcatccaatbtgctatgt | 2334 |
| OY | 1921 | ttttgtatagtttgcagttgcagtgttggtttgcatagagtggtgacatcgtgtatattctaaag    | 1980 |
| Db | 2335 | ttttgtatagtttgcagttgcagtgttggtttgcatagagtggtgacatcgtgtatattctaaag    | 2394 |
| OY | 1981 | catgttggaacacagatgaactctctctgtctctctctctctctctctctctctctctctcaacaaca | 2040 |
| Db | 2395 | catgttggaacacagatgaactctctctgtctctctctctctctctctctctctctctctcaacaaca | 2454 |
| OY | 2041 | aatgtgtcatatgaagacacatctccatctccatctcagtagagaagaactgtcttcatgtc       | 2100 |

|||||  
Db 2455 aatggtctatgaagacaactccacctatccattccattccaggaaactgltcttcacgtc 2514  
OY 2101 gatggaaaaacccaggtctatggaattctgggggtccacaactccagacttcggaaacagg 2160  
Db 2515 gatggaaaaacccaggtctatggaattctgggggtccacaactccagacttcggaaacagg 2514  
OY 2161 catgaccgctctactgaagggtctctagttgtgacaagaagaactggtgattatctacggaga 2220  
Db 2575 catgaccgctctactgaagggtctctagttgtgacaagaagaactggtgattatctacggaga 2614  
OY 2221 cagttatgaagataatttcagacatactgtcgtagttaaacaatgccaatgaaacaaagg 2280  
Db 2635 cagttatgaagataatttcagacatactgtcgtagttaaacaatgccaatgaaacaaagg 2694  
OY 2281 ctctcccccgaattccaagacaccctagcactagcgaagaacaaatttaatgcccaccac 2340  
Db 2695 ctctcccccgaat-----ccacc 2712  
OY 2341 agtcttgaaacgcatcaacggaataaactcgttactactcttcagtcagatcaagaga 2400  
Db 2713 agtcttgaaacgcatcaacggaataaactcgttactactcttcagtcagatcaagaga 2772  
OY 2401 aattgacctatgatataccaatatactcaagttgaatggaaggaagattttacaatttga 2460  
Db 2773 aattgacctatgatataccaatatactcaagttgaatggaaggaagattttacaatttga 2832  
OY 2461 tgaagatgaagaatcagaagcccccagcgttccaagaagaacacgaacattttatgtc 2520  
Db 2833 tgaagatgaagaatcagaagcccccagcgttccaagaagaacacgaacattttatgtc 2892  
OY 2521 tgcagtggaagagcgtcttgsgatatatggaatgaatggtccccaatgtctctaagaacacag 2580  
Db 2893 tgcagtggaagagcgtcttgsgatatatggaatgaatggtccccaatgtctctaagaacacag 2952  
OY 2581 ggctcagagtgagtggtgtccctcaagtcgaagaagtggtttccaggaatttactgtg 2640  
Db 2953 ggctcagagtgagtggtgtccctcaagtcgaagaagtggtttccaggaatttactgtg 3012  
OY 2641 ccccttactcagcccttaacacgttgagaagaatgaagaactttggagctccctggggcc 2700  
Db 3013 ccccttactcagcccttaacacgttgagaagaatgaagaactttggagctccctggggcc 3072  
OY 2701 atataaagaagcagaagttgaagaataatcatcgttaacttcagaacatcaaggcctctcg 2760  
Db 3073 atataaagaagcagaagttgaagaataatcatcgttaacttcagaacatcaaggcctctcg 3132  
OY 2761 tccctattctcttatcttactagccttattcttatggagaagaatcagaaggcgaagagaga 2820  
Db 3133 tccctattctcttatcttactagccttattcttatggagaagaatcagaaggcgaagagaga 3192  
OY 2821 accctagaaaaaacttcttcaagccttaatgaacacaaacttactttggaaagtgcacaa 2880  
Db 3193 accctagaaaaaacttcttcaagccttaatgaacacaaacttactttggaaagtgcacaa 3252  
OY 2881 tcaatatgacacccaactaaagaatgagttgacgtgcaaaagcctgggctattctctgtgt 2940  
Db 3253 tcaatatgacacccaactaaagaatgagttgacgtgcaaaagcctgggctattctctgtgt 3312  
OY 2941 tgaacctggaaaaaagatgtgtcactcagaagcctgatatggaccctctgtctgccaactaa 3000  
Db 3313 tgaacctggaaaaaagatgtgtcactcagaagcctgatatggaccctctgtctgccaactaa 3372  
OY 3001 cagactgaacccctgtcctcattgggagacaaagtgcacagtaacagaatttgcgtgttttcac 3060  
Db 3373 cagactgaacccctgtcctcattgggagacaaagtgcacagtaacagaatttgcgtgttttcac 3432  
OY 3061 catctcttgaatgaacacaaagctgtgtaactcacttgaataataatggaagaacactgacggc 3120  
Db 3433 catctcttgaatgaacacaaagctgtgtaactcacttgaataataatggaagaacactgacggc 3492  
OY 3121 tccctcgaataatccagatggaagaatcccaactttaaagagaattatcgcttccatgcaat 3180  
|||||

Db 3493 tccctcgaataatccagatggaagaatcccaactttaaagagaattatcgcttccatgcaat 3552  
OY 3181 caatggtctacataatgatgatactactcgtggttagtaagaatggtccagaatcaagaattcg 3240  
Db 3553 caatggtctacataatgatgatactactcgtggttagtaagaatggtccagaatcaagaattcg 3612  
OY 3241 atggtatctgtcgaacatggtgcagcaatgtaaaacatccatctcatcttcaattgagaca 3300  
Db 3613 atggtatctgtcgaacatggtgcagcaatgtaaaacatccatctcatcttcaattgagaca 3672  
OY 3301 tgtgttactcgttgcgaaaaaaagagagataaataatggaacgttaacatcttaccgg 3360  
Db 3673 tgtgttactcgttgcgaaaaaaagagagataaataatggaacgttacaactcttaccgg 3732  
OY 3361 tgtttttgacacagtggaatgtttacatccaaagcttggaatttggcgggtggaattgccc 3420  
Db 3733 tgtttttgacacagtggaatgtttacatccaaagcttggaatttggcgggtggaattgccc 3792  
OY 3421 tattggcgaacatctacatgctggtggaatgacacatttctctgtgttaccagaataatgtg 3480  
Db 3793 tattggcgaacatctacatgctggtggaatgacacatttctctgtgttaccagaataatgtg 3852  
OY 3481 tcaagctccctctgggaatggtctcttgacacatataagatttcaagataagcttcagg 3540  
Db 3853 tcaagctccctctgggaatggtctcttgacacatataagatttcaagataagcttcagg 3912  
OY 3541 acatatggaacagtggtgcccccaagcttggtccagactctatctccggaatcaatcagc 3600  
Db 3913 acatatggaacagtggtgcccccaagcttggtccagactctatctccggaatcaatcagc 3972  
OY 3601 ctggagacacaaaggaccccttctctggaatcaaggtggaatctgttggccaacaaatgatt 3660  
Db 3973 ctggagacacaaaggaccccttctctggaatcaaggtggaatctgttggccaacaaatgatt 4032  
OY 3661 tcaagcatcaaaacccaagggtgcccgtcagaagaattctccagcttataatcttcagtt 3720  
Db 4033 tcaagcatcaaaacccaagggtgcccgtcagaagaattctccagcttataatcttcagtt 4092  
OY 3721 tatcatcatgtatagcttcttatgagaagaatgtgcagacttactcagaagaattccagtg 3780  
Db 4093 tatcatcatgtatagcttcttatgagaagaatgtgcagacttactcagaagaattccagtg 4152  
OY 3781 aaacctaaatggtctctcttggcaatgtggaatcatcttgggaataaaacacaaatatttaa 3840  
Db 4153 aaacctaaatggtctctcttggcaatgtggaatcatcttgggaataaaacacaaatatttaa 4212  
OY 3841 cccctcaatatgtgtcgtgatacatccgttgcacccaactcaatataagatctgcagac 3900  
Db 4213 cccctcaatatgtgtcgtgatacatccgttgcacccaactcaatataagatctgcagac 4272  
OY 3901 tcttcgcatggaatgtgaggtcgtgatttaataatgtgacgaatgccaatgggaattgga 3960  
Db 4273 tcttcgcatggaatgtgaggtcgtgatttaataatgtgacgaatgccaatgggaattgga 4332  
OY 3961 gagtaagaacatatcagaatgacacagattactgtcttaacttaacaaatagtcttcg 4020  
Db 4333 gagtaagaacatatcagaatgacacagattactgtcttaacttaacaaatagtcttcg 4392  
OY 4021 caactgtctccttcaaaaagctgacttcaacctcaagggaaggaatggtcttgaagacc 4080  
Db 4393 caactgtctccttcaaaaagctgacttcaacctcaagggaaggaatggtcttgaagacc 4452  
OY 4081 tcaagtgaaatccaaaagaatggtctgcaagtgcgttccagaagaacaaatgaagaatcac 4140  
Db 4453 tcaagtgaaatccaaaagaatggtctgcaagtgcgttccagaagaacaaatgaagaatcac 4512  
OY 4141 aggaatgaactactcaggaggaataaatctcgtcttaccagaatgtatgtgaaggagttcct 4200  
Db 4513 aggaatgaactactcaggaggaataaatctcgtcttaccagaatgtatgtgaaggagttcct 4572  
OY 4201 catctccagacagtcagaatggtccatcgaatggaatctcttttccgaatggtgcaagtaaa 4260  
Db 4573 catctccagacagtcagaatggtccatcgaatggaatctcttttccgaatggtgcaagtaaa 4632  
|||||





QY 1501 tgaatgcgtcccttctgtatctcaaggagattaccaaaggtgttaaacatttgaagatctt 1560  
| | | | |  
DB 1915 tgaatgcgtcccttctgtatctcaaggagattaccaaaggtgttaaacatttgaagatctt 1974  
| | | | |  
QY 1561 tccaattcttcggggaataattccaataataatgagagtgactgtgagaagatgtggcc 1620  
| | | | |  
DB 1975 tccaattcttcggggaataataatccaataataatgagagtgactgtgagaagatgtggcc 2034  
| | | | |  
QY 1621 aactaaatcagatccctgtgtccctgcagccgtctactctagtctgttaatatgagaag 1680  
| | | | |  
DB 2035 aactaaatcagatccctgtgtccctgcagccgtctactctagtctgttaatatgagaag 2094  
| | | | |  
QY 1681 agatctagcttcagagactcatctggccctctccatctgtctacaagaatctgtatgaca 1740  
| | | | |  
DB 2095 agatctagcttcagagactcatctggccctctccatctgtctacaagaatctgtatgaca 2154  
| | | | |  
QY 1741 aagaagaaacagataatgtgcagacaagagaaatgtatccctgtttctgtatattgata 1800  
| | | | |  
DB 2155 aagaagaaacagataatgtgcagacaagagaaatgtatccctgtttctgtatattgata 2214  
| | | | |  
QY 1801 gaaccgaagctgtgtacctcacaagaaatatacaacgcttctcccaatccagctgtgagt 1860  
| | | | |  
DB 2215 gaaccgaagctgtgtacctcacaagaaatatacaacgcttctcccaatccagctgtgagt 2274  
| | | | |  
QY 1861 gcaagcttgagatccagagatccaaagcctccaacatcatgacagatcaatgtgatatgt 1920  
| | | | |  
DB 2275 gcaagcttgagatccagagatccaaagcctccaacatcatgacagatcaatgtgatatgt 2334  
| | | | |  
QY 1921 ttgttgatagtttcagatgtgcagttgtttgtgacagagtgagacatctgtatcttaag 1980  
| | | | |  
DB 2335 ttgttgatagtttcagatgtgcagttgtttgtgacagagtgagacatctgtatcttaag 2394  
| | | | |  
QY 1981 catgtgagcacaagactgtactctctgtctctctctctctctctctctctctctctctct 2040  
| | | | |  
DB 2395 catgtgagcacaagactgtactctctgtctctctctctctctctctctctctctctctct 2454  
| | | | |  
QY 2041 aatgtctatagaaagacaactacacccatctccatctccagagaaactgtctcatgtc 2100  
| | | | |  
DB 2455 aatgtctatagaaagacaactacacccatctccatctccagagaaactgtctcatgtc 2514  
| | | | |  
QY 2101 gatggaaaaacccaggtgtactgtggtgtgcacaacacagactcttgagaacagag 2160  
| | | | |  
DB 2515 gatggaaaaacccaggtgtactgtggtgtgcacaacacagactcttgagaacagag 2574  
| | | | |  
QY 2161 catgacgcgccttaactgaaggtttctctgtgtgacaagaacactgtgatatctacgagga 2220  
| | | | |  
DB 2575 catgacgcgccttaactgaaggtttctctgtgtgacaagaacactgtgatatctacgagga 2634  
| | | | |  
QY 2221 cagttatgaagataattcagacatactgtcgtgagtaaaaaacaatgcatctgaacaaag 2280  
| | | | |  
DB 2635 cagttatgaagataattcagacatactgtcgtgagtaaaaaacaatgcatctgaacaaag 2694  
| | | | |  
QY 2281 ctctccccaagattcaagacacccctagcacttaggcaaaagcaatttaatgcccaccgc 2340  
| | | | |  
DB 2695 ctctccccaagattcaagacacccctagcacttaggcaaaagcaatttaatgcccaccgc 2712  
| | | | |  
QY 2341 agtcttgaaaaacgcacatcaacggygaataaactcgttactactctctcagtcagatcaagaga 2400  
| | | | |  
DB 2713 agtcttgaaaaacgcacatcaacggygaataaactcgttactactctctcagtcagatcaagaga 2772  
| | | | |  
QY 2401 aattgactatgtatgataccatcatcagttgtaaatggaagaagaagatttgaacttatga 2460  
| | | | |  
DB 2773 aattgactatgtatgataccatcatcagttgtaaatggaagaagaagatttgaacttatga 2832  
| | | | |  
QY 2461 tgaagatgaanaatcagagcccccgcagctctcaaaaagaagaacagaaactatttatgac 2520  
| | | | |  
DB 2833 tgaagatgaanaatcagagcccccgcagctctcaaaaagaagaacagaaactatttatgac 2892  
| | | | |  
QY 2521 tgcagtgagagagctcttgagattatgagatgagtagtccccaatgttcttaagaanaag 2580  
| | | | |  
DB 2893 tgcagtgagagagctcttgagattatgagatgagtagtccccaatgttcttaagaanaag 2952  
| | | | |  
QY 2581 ggcctcagagtgagtgatgctccctcagttcagaagaagttgtttctccaagaatttactgtg 2640  
| | | | |

DB 2953 ggcctcagtgagtgatgctccctcagttcagaagaagttgtttctccaagaatttactgtg 3012  
| | | | |  
QY 2641 ctctcttactcagcccttataccgttgagaaactaaatgaacatttggactcttgggccc 2700  
| | | | |  
DB 3013 ctctcttactcagcccttataccgttgagaaactaaatgaacatttggactcttgggccc 3072  
| | | | |  
QY 2701 atataaagcagaagttgaagataataatcatgttgaacttccagaataacagccctctg 2760  
| | | | |  
DB 3073 atataaagcagaagttgaagataataatcatgttgaacttccagaataacagccctctg 3132  
| | | | |  
QY 2761 tccctattcctctatctctgaccttatctctatgaagaatcagaagcagaagcaga 2820  
| | | | |  
DB 3133 tccctattcctctatctctgaccttatctctatgaagaatcagaagcagaagcaga 3192  
| | | | |  
QY 2821 acctagaanaaacttctcagacctaagtaaaccaaaacttactcttgaagaatgtaaaa 2880  
| | | | |  
DB 3193 acctagaanaaacttctcagacctaagtaaaccaaaacttactcttgaagaatgtaaaa 3252  
| | | | |  
QY 2881 tcatatgcaaccacataaagatgagttgtactgtcaaaagccttggtctatctctgtatgt 2940  
| | | | |  
DB 3253 tcatatgcaaccacataaagatgagttgtactgtcaaaagccttggtctatctctgtatgt 3312  
| | | | |  
QY 2941 tgaacttgaaaaaagatgtgcaactcaggtgtgagccctctgtgtgccaactaa 3000  
| | | | |  
DB 3313 tgaacttgaaaaaagatgtgcaactcaggtgtgagccctctgtgtgccaactaa 3372  
| | | | |  
QY 3001 cacactgaaccctctcactcttgagagacaggtgacagttacagaatcttgccctgttttcc 3060  
| | | | |  
DB 3373 cacactgaaccctctcactcttgagagacaggtgacagttacagaatcttgccctgttttcc 3432  
| | | | |  
QY 3061 catcttctgataagacccaagaagctgtgtactcactgtgaataatatagaaagaacatgcagggc 3120  
| | | | |  
DB 3433 catcttctgataagacccaagaagctgtgtactcactgtgaataatatagaaagaacatgcagggc 3492  
| | | | |  
QY 3121 tcccttgcaatatccagttggaagaatcccaacttttaagagaattatcgcttccatgcaat 3180  
| | | | |  
DB 3493 tcccttgcaatatccagttggaagaatcccaacttttaagagaattatcgcttccatgcaat 3552  
| | | | |  
QY 3181 caatggtctacataatgtgatacactacgtgtgtagtaatgaggtcagaatcaagaatgag 3240  
| | | | |  
DB 3553 caatggtctacataatgtgatacactacgtgtgtagtaatgaggtcagaatcaagaatgag 3612  
| | | | |  
QY 3241 atggtatctgtctcagcaatggtgcagcaatgaaacatccatcttatcatctcagtgagca 3300  
| | | | |  
DB 3613 atggtatctgtctcagcaatggtgcagcaatgaaacatccatcttatcatctcagtgagca 3672  
| | | | |  
QY 3301 tgtgttactgttactcagaaagaagagatataaatgtgcactgttacaatctctatccag 3360  
| | | | |  
DB 3673 tgtgttactgttactcagaaagaagagatataaatgtgcactgttacaatctctatccag 3732  
| | | | |  
QY 3361 tgtttttgagacagttgaaatgttaccatccaaagctggaatttggcggtggaatgacct 3420  
| | | | |  
DB 3733 tgtttttgagacagttgaaatgttaccatccaaagctggaatttggcggtggaatgacct 3792  
| | | | |  
QY 3421 tatgtgcagatctacatagctgtgagatgagacacttttctgtgtgtatagaacaataatgt 3480  
| | | | |  
DB 3793 tatgtgcagatctacatagctgtgagatgagacacttttctgtgtgtatagaacaataatgt 3852  
| | | | |  
QY 3481 tcaagactcccccgtggaatgtgctctcgtgacacatataagaatttccagataacagcttcag 3540  
| | | | |  
DB 3853 tcaagactcccccgtggaatgtgctctcgtgacacatataagaatttccagataacagcttcag 3912  
| | | | |  
QY 3541 acaatatgagacagttggcccccaagctgtgcagaactctatctcccgagatcaatcaatgc 3600  
| | | | |  
DB 3913 acaatatgagacagttggcccccaagctgtgcagaactctatctcccgagatcaatcaatgc 3972  
| | | | |  
QY 3601 ctggaagaccaaagagcccttctcttgatcaaggtgtgactgtgttgacccaatgatat 3660  
| | | | |  
DB 3973 ctggaagaccaaagagcccttctcttgatcaaggtgtgactgtgtgttgacccaatgatat 4032  
| | | | |  
QY 3661 tcaagacataaagaccagaggtgcccgtcagaagttctccaagctctacatctctcagtt 3720  
| | | | |



```
Db 4033 tcacgacatcaagaccagggtgcccgtccagaaagttctccagcctctacatctcagtt 4092
QY 3721 tatcatcatgtatagcttgaatggaaagatggtgcagactatcagggaattccactg 3780
Db 4093 tatcaatcaatgataagctctgaagggaagatggtgcagactatcagggaattccactg 4152
QY 3781 aaccttaatgtctctctcttggcaatgtgatatcatctggataaacaacatacttttaa 3840
Db 4153 aaccttaatgtctctctcttggcaatgtgatatcatctggataaacaacatacttttaa 4212
QY 3841 cccctcaatattgtctcgaacacacgttgcacccaactcattatagcattcgcagac 3900
Db 4213 cccctcaatattgtctcgaacacacgttgcacccaactcattatagcattcgcagac 4272
QY 3901 tcttcgcatggaatgatatggtgctgtatataaattgagatgccatctgggaatgga 3960
Db 4273 tcttcgcatggaatgatatggtgctgtatataaattgagatgccatctgggaatgga 4332
QY 3961 gagtaagacataatcagatgacagattactgtctacatcttaccataatgtttgc 4020
Db 4333 gagtaagacataatcagatgacagattactgtctacatcttaccataatgtttgc 4392
QY 4021 caccctgtctcctccaaagctcgtactcaactccaaagggaagataatgcttgagacc 4080
Db 4393 caccctgtctcctccaaagctcgtactcaactccaaagggaagataatgcttgagacc 4452
QY 4081 tcaggtgaataatccaaagatggtgtgaagtggaactccaaagacataatgaatgaac 4140
Db 4453 tcaggtgaataatccaaagatggtgtgaagtggaactccaaagacataatgaatgaac 4512
QY 4141 aggaagtaactactcagaaggatlaaatctctgtctaccagatgatagtgaagagttcct 4200
Db 4513 aggaagtaactactcagaaggatlaaatctctgtctaccagatgatagtgaagagttcct 4572
QY 4201 catctccagcagatcaagatggtccatcagtgagactcctttttcagaatggtcagaataa 4260
Db 4573 catctccagcagatcaagatggtccatcagtgagactcctttttcagaatggtcagaataa 4632
QY 4261 ggttttcagggaatcaagactccttcaactctgtgtggaactcctctgaccaccgctt 4320
Db 4633 ggttttcagggaatcaagactccttcaactctgtgtggaactcctctgaccaccgctt 4692
QY 4321 acgagactcgtactcctcgaattaccccccaagatgtgtgacacagatggtccctgagat 4380
Db 4693 acgagactcgtactcctcgaattaccccccaagatgtgtgacacagatggtccctgagat 4752
QY 4381 ggaagttctggtgctgcgagacagagactctactga 4417
Db 4753 ggaagttctggtgctgcgagacagagactctactga 4789

RESULT 6
US-09-470-618-13
; Sequence 13, Application US/09470618
; Patent No. 6200560
; GENERAL INFORMATION:
; APPLICANT: Couto, Linda B.
; APPLICANT: Colosi, Peter C.
; TITLE OF INVENTION: Adeno-Associated Vectors for Expression of Factor VIII
; FILE REFERENCE: by Target Cells
; CURRENT APPLICATION NUMBER: US/09/470,618
; EARLIER APPLICATION NUMBER: 09/364,862
; EARLIER FILING DATE: 1999-07-30
; EARLIER APPLICATION NUMBER: 60/125,974
; EARLIER FILING DATE: 1999-03-24
; EARLIER APPLICATION NUMBER: 60/104,994
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 11933
```

```
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-470-618-13

Query Match
Best Local Similarity 98.2%; Score 4249.4; DB 4; Length 11933;
Matches 4338; Conservative 0; Mismatches 1; Indels 78; Gaps 1;

QY 1 catgcaatagagctctccacgtctctctctctctctctctctctctctctctctctctctct 60
Db 426 catgcaatagagctctccacgtctctctctctctctctctctctctctctctctctctctct 485
QY 61 caccagaataactaccctggtgagtgagtggaactgtcaatgaggaactatagcaaatgagatct 120
Db 486 caccagaataactaccctggtgagtgagtggaactgtcaatgaggaactatagcaaatgagatct 545
QY 121 cgtgtgagctggtcgtgtgagcgaagatttctcctagagtgccaaatctcttccattcaa 180
Db 546 cgtgtgagctggtcgtgtgagcgaagatttctcctagagtgccaaatctcttccattcaa 605
QY 181 caccctagtcgtgtacaaaagactctgtgtgtgagattcacggttcacatttcaaat 240
Db 606 caccctagtcgtgtacaaaagactctgtgtgtgagattcacggttcacatttcaaat 665
QY 241 cgtctaaagcgaagccacccctgagtggtgtctgtaggtctccatccacacagctgagttta 300
Db 666 cgtctaaagcgaagccacccctgagtggtgtctgtaggtctccatccacacagctgagttta 725
QY 301 tgataagtgatcatatcaactaaagacatgtgtccctccctcgtgagcttcaatctgt 360
Db 726 tgataagtgatcatatcaactaaagacatgtgtccctccctcgtgagcttcaatctgt 785
QY 361 tgggtatctactctggaagagctctctgagtgagtgagtgagtgagtgagtgagtgagtgagtg 420
Db 786 tgggtatctactctggaagagctctctgagtgagtgagtgagtgagtgagtgagtgagtgagtg 845
QY 421 ggaagaaagaaatgataaagctctcctctggtgtggaagcaataatgcttggaagctct 480
Db 846 ggaagaaagaaatgataaagctctcctctggtgtggaagcaataatgcttggaagctct 905
QY 481 gaaagaaatgtgtcgaatgtggtcctgtgacacactgtgtcctactactatctactctctca 540
Db 906 gaaagaaatgtgtcgaatgtggtcctgtgacacactgtgtcctactactatctactctctca 965
QY 541 tgtgagcctgtgtaaaagactgaattcaggtcctcatctgagccctactagtatgtagaga 600
Db 966 tgtgagcctgtgtaaaagactgaattcaggtcctcatctgagccctactagtatgtagaga 1025
QY 601 agggagctctgccaagaaagaaagacacagacctgtgcacaaattatactactcttggct 660
Db 1026 agggagctctgccaagaaagaaagacacagacctgtgcacaaattatactactcttggct 1085
QY 661 atttgatgaagggaagatgtgtgacacagaagaaactcctgtatgcagatagaga 720
Db 1086 atttgatgaagggaagatgtgtgacacagaagaaactcctgtatgcagatagaga 1145
QY 721 tctgtcatctgtctggtcgtggtccttaaaatgacacagcgaatgtgtatgtaaacagctc 780
Db 1146 tctgtcatctgtctggtcgtggtccttaaaatgacacagcgaatgtgtatgtaaacagctc 1205
QY 781 tctgtcaggtctgtatgtatggtccacaggaatcagctatgttgcatgtatgtgaatgag 840
Db 1206 tctgtcaggtctgtatgtatggtccacaggaatcagctatgttgcatgtatgtgaatgag 1265
QY 841 caccactcctgagtgacatcaatattcctggaagtgacacatctctgtgtgaagaaca 900
Db 1266 caccactcctgagtgacatcaatattcctggaagtgacacatctctgtgtgaagaaca 1325
QY 901 tcgcagagctccttggaatctgcgcaataacttctctactgtcgaacactctgtat 960
```

|   |      |   |      |
|---|------|---|------|
| D | 1326 | tcgcacgagctccttgcgaaactctgcacaaatacttcccttacttgcctcaaaacatcttgc | 1385 |
| Q | 961  | ggaaccttgacagcttctactgtttgtatcatatctctccacacacatgatagtgcata     | 1020 |
| D | 1386 | ggaaccttgacagcttctactgtttgtatcatatctctccacacacatgatagtgcata     | 1445 |
| Q | 1021 | agcttatgtcaaaagtacagacgtctgcagaggaaccccaactacgaatgaaataatga     | 1080 |
| D | 1446 | agcttatgtcaaaagtacagacgtctgcagaggaaccccaactacgaatgaaataatga     | 1505 |
| Q | 1081 | agagagcggaagactatgatgatgatcttactgtatcttgcgaatggatgtgcagattga    | 1140 |
| D | 1506 | agagagcggaagactatgatgatgatcttactgtatcttgcgaatggatgtgcagattga    | 1565 |
| Q | 1141 | tgaatgaacatctccctcccttaccacatctgcctcagcttgcacagaaagcaatccaaac   | 1200 |
| D | 1566 | tgaatgaacatctccctcccttaccacatctgcctcagcttgcacagaaagcaatccaaac   | 1625 |
| Q | 1201 | ttgggtacatatacatctgcctgcgaagagagagactggactatgctcccttagtctcgc    | 1260 |
| D | 1626 | ttgggtacatatacatctgcctgcgaagagagagactggactatgctcccttagtctcgc    | 1685 |
| Q | 1261 | ccccgctgcacagaagtataaaagtcacataatttgaacaatggccctcaagcgattgtag   | 1320 |
| D | 1686 | ccccgctgcacagaagtataaaagtcacataatttgaacaatggccctcaagcgattgtag   | 1745 |
| Q | 1321 | gaagtaacaaaagctccgactttagcatcacacagatgaaacctttaagactctgtagac    | 1380 |
| D | 1746 | gaagtaacaaaagctccgactttagcatcacacagatgaaacctttaagactctgtagac    | 1805 |
| Q | 1381 | tattcagcatgaatcaggaatcttgcgaccttacttattgcgggaagltgcgaacacact    | 1440 |
| D | 1806 | tattcagcatgaatcaggaatcttgcgaccttacttattgcgggaagltgcgaacacact    | 1865 |
| Q | 1441 | gtttgatataattaaagaatacaagcaagcacacataataatctcacccctcaacgaatacac | 1500 |
| D | 1866 | gtttgatataattaaagaatacaagcaagcacacataataatctcacccctcaacgaatacac | 1925 |
| Q | 1501 | tgaatgcctgcctcttctatccaaggaatataccaagaagtgtaaaacatctgaaagattt   | 1560 |
| D | 1926 | tgaatgcctgcctcttctatccaaggaatataccaagaagtgtaaaacatctgaaagattt   | 1985 |
| Q | 1561 | tcacaattcgcagagagaataatatacaaaataatgacagctgcataagaagtgcggcc     | 1620 |
| D | 1986 | tcacaattcgcagagagaataatatacaaaataatgacagctgcataagaagtgcggcc     | 2045 |
| Q | 1621 | aaactaaatacagatccctgcggtgcagccgctataactcagttctgttaataatgcggag   | 1680 |
| D | 2046 | aaactaaatacagatccctgcggtgcagccgctataactcagttctgttaataatgcggag   | 2105 |
| Q | 1681 | agatctagacttcagagactcatctgcctctccatcatctgcctacaaaagaatctgtatga  | 1740 |
| D | 2106 | agatctagacttcagagactcatctgcctctccatcatctgcctacaaaagaatctgtatga  | 2165 |
| Q | 1741 | aagaggaagaacagataatgtccagacaaaggaatgtcatccgtttctctgtatattgata   | 1800 |
| D | 2166 | aagaggaagaacagataatgtccagacaaaggaatgtcatccgtttctctgtatattgata   | 2225 |
| Q | 1801 | gaacccaagaactgttactctcagagaaatatacaacgctttctccccaatccagcttgagt  | 1860 |
| D | 2226 | gaacccaagaactgttactctcagagaaatatacaacgctttctccccaatccagcttgagt  | 2285 |
| Q | 1861 | gaagcttgaggaatccagagatttccaagctccaacatcatgcacagataatgtctatgt    | 1920 |
| D | 2286 | gaagcttgaggaatccagagatttccaagctccaacatcatgcacagataatgtctatgt    | 2345 |
| Q | 1921 | ctttgataatctgcagcttgcagcttgcgttcgtgcataaggtgcgacatactgtatctaa   | 1980 |
| D | 2346 | ctttgataatctgcagcttgcagcttgcgttcgtgcataaggtgcgacatactgtatctaa   | 2405 |
| Q | 1981 | catcttgagacacagactgactctcttctgtctctctctctctctctctctctctctctct   | 2040 |
| D | 2406 | catcttgagacacagactgactctcttctgtctctctctctctctctctctctctctctct   | 2465 |

---

|   |      |  |      |
|---|------|--|------|
| Q | 2041 | aatgtctatgaagaacacacactcacaccttaccattccattctcagagaaactgtctcatgc  | 2100 |
| D | 2466 | aatgtctatgaagaacacacactcacaccttaccattccattctcagagaaactgtctcatgc  | 2525 |
| Q | 2101 | gatbgaaaacccaggtcttatgattctgcgggtgcacaaactcgaactcttcggaacagag    | 2160 |
| D | 2526 | gatbgaaaacccaggtcttatgattctgcgggtgcacaaactcgaactcttcggaacagag    | 2585 |
| Q | 2161 | catgaccccttacttgaggtttctatgttgcacaaagaacactggtgatattatcagaaga    | 2220 |
| D | 2586 | catgaccccttacttgaggtttctatgttgcacaaagaacactggtgatattatcagaaga    | 2645 |
| Q | 2221 | caattatgaagaattatcagacatactgtcgtgagtaaaacaatgcatgtaaccaagaag     | 2280 |
| D | 2646 | caattatgaagaattatcagacatactgtcgtgagtaaaacaatgcatgtaaccaagaag     | 2705 |
| Q | 2281 | cttctccagaagaattcaagacacacctgacataagcaaaagcaatttaatgccaacccacc   | 2340 |
| D | 2706 | cttctccagaagaattcaagacacacctgacataagcaaaagcaatttaatgccaacccacc   | 2709 |
| Q | 2341 | agcttgaagaagccatcaacgaggaataactcgtactactcttcagtcagatcaagaaga     | 2400 |
| D | 2710 | agcttgaagaagccatcaacgaggaataactcgtactactcttcagtcagatcaagaaga     | 2747 |
| Q | 2401 | aattgactatgatgatataccatatacagttgaaatgaaagaagaagatttgcattatga     | 2460 |
| D | 2748 | aattgactatgatgatataccatatacagttgaaatgaaagaagaagatttgcattatga     | 2807 |
| Q | 2461 | tgaagatgaaaaatcagaagcccccgcagacttcaaaagaacacagacactatttatgc      | 2520 |
| D | 2808 | tgaagatgaaaaatcagaagcccccgcagacttcaaaagaacacagacactatttatgc      | 2867 |
| Q | 2521 | tgacagtgagagagctctgcggattatagagatgagtaagcttccccaacatgcttaagaacag | 2580 |
| D | 2868 | tgacagtgagagagctctgcggattatagagatgagtaagcttccccaacatgcttaagaacag | 2927 |
| Q | 2581 | ggctcagagctgcagctgtccctcagttcgaagaagaatgtgtttccagaagaatttactgag  | 2640 |
| D | 2928 | ggctcagagctgcagctgtccctcagttcgaagaagaatgtgtttccagaagaatttactgag  | 2987 |
| Q | 2641 | ctccttactcagcccttatatacgttgcagaaactaaatgaaacatttgcgctccctgcggcc  | 2700 |
| D | 2988 | ctccttactcagcccttatatacgttgcagaaactaaatgaaacatttgcgctccctgcggcc  | 3047 |
| Q | 2701 | atatataagagcagaagttgaagaataatcatcagtgtaacttccaagaatccagccctcgc   | 2760 |
| D | 3048 | atatataagagcagaagttgaagaataatcatcagtgtaacttccaagaatccagccctcgc   | 3107 |
| Q | 2761 | tcacctatccttatactctagccttattcttataaggaagatcagaagcaagagacaga      | 2820 |
| D | 3108 | tcacctatccttatactctagccttattcttataaggaagatcagaagcaagagacaga      | 3167 |
| Q | 2821 | acctagaanaaactttgtcaagcctaaatgaacaaaacttatttttgaagaagtgaacaa     | 2880 |
| D | 3168 | acctagaanaaactttgtcaagcctaaatgaacaaaacttatttttgaagaagtgaacaa     | 3227 |
| Q | 2881 | tcatatgcaccccaataaagatgatttgcagtcaaaagcctgcgctatttctctatgt       | 2940 |
| D | 3228 | tcatatgcaccccaataaagatgatttgcagtcaaaagcctgcgctatttctctatgt       | 3287 |
| Q | 2941 | tgaacttgaaaagaatgttgcactcagagcctgtattgaaacccctctgcgtgcacacata    | 3000 |
| D | 3288 | tgaacttgaaaagaatgttgcactcagagcctgtattgaaacccctctgcgtgcacacata    | 3347 |
| Q | 3001 | cacactgaacccctgcctcatatgcagacaaagtgcacagtaacagaatttgcctgtttttcac | 3060 |
| D | 3348 | cacactgaacccctgcctcatatgcagacaaagtgcacagtaacagaatttgcctgtttttcac | 3407 |
| Q | 3061 | catctttgatgagacaaagaagctgcgtactcacttgaaataatgaaagaagaacatgcagggc | 3120 |
| D | 3408 | catctttgatgagacaaagaagctgcgtactcacttgaaataatgaaagaagaacatgcagggc | 3467 |

QY 3121 tccctgcaatataccaga tggagaagatcccaactttaagaagaattatcgcttcacatgcaat 3180  
|||||  
Db 3468 tccctgcaatataccaga tggagaagatcccaactttaagaagaattatcgcttcacatgcaat 3527  
QY 3181 caatgctcaataatgataatacaactacgtgcttaagtaatggtccaga tcaaaagatcgc 3240  
|||||  
Db 3528 caatgctcaataatgataatacaactacgtgcttaagtaatggtccaga tcaaaagatcgc 3587  
QY 3241 atggtatcgtctcaagatgggcaagcaatgaaacacatcattatatttcacgttcgagaca 3300  
Db 3588 atggtatcgtctcaagatgggcaagcaatgaaacacatcattatatttcacgttcgagaca 3647  
QY 3301 tgtgttcaactgatacaaaaaaagggagataaaatgycactgtacacatctacacagg 3360  
3648 tgtgttcaactgatacaaaaaaagggagataaaatgycactgtacacatctacacagg 3707  
QY 3361 tgtttttgagaacagtggaatgttaacacatccaaagcttgaaattgycgggtggaatgcct 3420  
Db 3708 tgtttttgagaacagtggaatgttaacacatccaaagcttgaaattgycgggtggaatgcct 3767  
QY 3421 tattggcgagacattacatgtcttgagatgacacattttctgtgtgacagcaataatgt 3480  
Db 3768 tattggcgagacattacatgtcttgagatgacacattttctgtgtgacagcaataatgt 3827  
QY 3481 tcaagactcccttggaatgagctctctgacacattagagatttcacagatlcacagcttcagg 3540  
Db 3828 tcaagactcccttggaatgagctctctgacacattagagatttcacagatlcacagcttcagg 3887  
QY 3541 acaatatgagacagtgggccccaagctgycagacattcattatcccgatccaatcaatgc 3600  
Db 3888 acaatatgagacagtgggccccaagctgycagacattcattatcccgatccaatcaatgc 3947  
QY 3601 ctggagacacaaagagccctttcttgatcagaagtgagatcctcttgacacaaatgtatat 3660  
Db 3948 ctggagacacaaagagccctttcttgatcagaagtgagatcctcttgacacaaatgtatat 4007  
QY 3661 tcacgagcatcaaaagacccaggtgcccgtcagagaagttctccagcctcacaactcctcagtc 3720  
Db 4008 tcacgagcatcaaaagacccaggtgcccgtcagagaagttctccagcctcacaactcctcagtc 4067  
QY 3721 tatcatcatgtatagctcttgaatgggaagaagtgycagacattatccgaaatccacatgcg 3780  
Db 4068 tatcatcatgtatagctcttgaatgggaagaagtgycagacattatccgaaatccacatgcg 4127  
QY 3781 aaacttaatgtcttctcttgacaaatgtagattatcctcgttgataaaacaaatctttttaa 3840  
Db 4128 aaacttaatgtcttctcttgacaaatgtagattatcctcgttgataaaacaaatctttttaa 4187  
QY 3841 cccctcaaatatgtctcgatatacatccgtttgacacccaacatcatatagcatctgcagcac 3900  
Db 4188 cccctcaaatatgtctcgatatacatccgtttgacacccaacatcatatagcatctgcagcac 4247  
QY 3901 tctctgcaatgagatgtgagtgctgtgattaaatgttcagcatgccaatctgggaatgga 3960  
Db 4248 tctctgcaatgagatgtgagtgctgtgattaaatgttcagcatgccaatctgggaatgga 4307  
QY 3961 gaaataagaacatatacagatcacagattactgttactactacttaacaaatgtttgc 4020  
Db 4308 gaaataagaacatatacagatcacagattactgttactactacttaacaaatgtttgc 4367  
QY 4021 caactgtgtccttcaaaaagctcgacttcaaccctcaagggaggaatgaatccctgcgagacc 4080  
Db 4368 caactgtgtccttcaaaaagctcgacttcaaccctcaagggaggaatgaatccctgcgagacc 4427  
QY 4081 tccagtggaataatcccaaaagatggtctgcaagtggacatcccaagaacaaatgaaatgcac 4140  
Db 4428 tccagtggaataatcccaaaagatggtctgcaagtggacatcccaagaacaaatgaaatgcac 4487  
QY 4141 aggaatgaacactcaagggaataaactctgcttaccacagatgtatgtgagaagatgcct 4200  
Db 4488 aggaatgaacactcaagggaataaactctgcttaccacagatgtatgtgagaagatgcct 4547  
QY 4201 catctccagcagtcacagatgagcatcagtgactctcttcttcagaaatgcaaaatgaa 4260

|||||  
Db 4548 catctccagcagtcacagatgagcatcagtgactctcttcttcagaaatgcaaaatgaa 4607  
QY 4261 ggttttcagggaaatcaaaactctcacaactctgagtgagtaactctctagaccacgcgt 4320  
Db 4608 ggttttcagggaaatcaaaactctcacaactctgagtgagtaactctctagaccacgcgt 4667  
QY 4321 actgactcgtactcttcagaaatcaaccccaagatgggtgcacacagatgccttgagat 4380  
Db 4668 actgactcgtactcttcagaaatcaaccccaagatgggtgcacacagatgccttgagat 4727  
QY 4381 ggaagtctggtgcgagggcacaggaagcctactga 4417  
Db 4728 ggaagtctggtgcgagggcacaggaagcctactga 4764

RESULT 7  
US-09-364-862-13  
; Sequence 6137 Application US/09364862  
; Patent No. 6221349  
; GENERAL INFORMATION:  
; APPLICANT: Couto, Linda B.  
; APPLICANT: Colosi, Peter C.  
; TITLE OF INVENTION: ADEMO-ASSOCIATED VECTORS FOR EXPRESSION OF FACTOR VIII  
; TITLE OF INVENTION: BY TARGET  
; TITLE OF INVENTION: CELLS  
; FILE REFERENCE: AVIGEN-03743  
; CURRENT APPLICATION NUMBER: US/09/364, 862  
; CURRENT FILING DATE: 1999-07-30  
; EARLIER APPLICATION NUMBER: 60/125, 974  
; EARLIER FILING DATE: 1999-03-24  
; EARLIER APPLICATION NUMBER: 60/104, 994  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 11933  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-364-862-13

Query Match 96.2%; Score 4249.4; DB 4; Length 11933;  
Best Local Similarity 98.2%; Pred. No. 0;  
Matches 4338; Conservative 0; Mismatches 1; Indels 78; Gaps 1;  
QY 1 catgcaaatagagctccacccgtctctctctgctgctcttgagattcgtcttaagtc 60  
Db 426 catgcaaatagagctccacccgtctctctctgctgctcttgagattcgtcttaagtc 485  
QY 61 caccagaagaatactactcctgggtgcagtggaactgtcatggaactatataatgcaaatgatatc 120  
Db 486 caccagaagaatactactcctgggtgcagtggaactgtcatggaactatataatgcaaatgatatc 545  
QY 121 cgggtggtgcctcgttgagacgaagatttctcctctaagtgagcacaatctttccattcaa 180  
Db 486 caccagaagaatactactcctgggtgcagtggaactgtcatggaactatataatgcaaatgatatc 545  
QY 546 cgggtggtgcctcgttgagacgaagatttctcctctaagtgagcacaatctttccattcaa 605  
Db 546 cgggtggtgcctcgttgagacgaagatttctcctctaagtgagcacaatctttccattcaa 605  
QY 181 caactagtcgtgtacaaaaagactcgtttgtgagaatcaacggttcacacttccaacat 240  
Db 606 caactagtcgtgtacaaaaagactcgtttgtgagaatcaacggttcacacttccaacat 665  
QY 241 cgttaagccaagggcaccctgagtggtctgctcaggtcctaaccatccaggtgaggttta 300  
Db 666 cgttaagccaagggcaccctgagtggtctgctcaggtcctaaccatccaggtgaggttta 725  
QY 301 tgatacagtggtatatacaacttaagaacatggttcccaactcgttaagcttcttatgtct 360  
Db 726 tgatacagtggtatatacaacttaagaacatggttcccaactcgttcaagcttcttatgtct 785  
QY 361 tgtgtatcctactgtaaaagcttctgagggagctgataatgatgatcagacagtcacaaag 420

```
|||||
Db 786 tggctgtatccctactggaagcctcttgaggagctggaatataatgatactgacacgtcaag 845
OY 421 ggggaaagaagatgataaagctctccctgggtggaagccatacatatgctctgcaagctct 480
Db 846 ggggaaagaagatgataaagctctccctgggtggaagccatacatatgctctgcaagctct 905
OY 481 gaaagaatggtccaaatggtcccttgaccccaatggtgcttaactactatcatctttctca 540
Db 906 gaaagaatggtccaaatggtcccttgaccccaatggtgcttaactactatcatctttctca 965
OY 541 tggggaactggttaaaagactggaatcgaagcctattggagccctactacgtatgtagaaga 600
Db 966 tggggaactggttaaaagactggaatcgaagcctattggagccctactacgtatgtagaaga 1025
OY 601 agggagctctggcccaaggaaaagacacagaccccttgacaaaattataactcttggcgt 660
Db 1026 agggagctctggcccaaggaaaagacacagaccccttgacaaaattataactcttggcgt 1085
OY 661 attctgataagggaaaagcttggtcacctcagaaaaagaactccttgatgacaggaagaaga 720
Db 1086 attctgataagggaaaagcttggtcacctcagaaaaagaactccttgatgacaggaagaaga 1145
OY 721 tggctgcatcgtcctgggctcggtctaaatgacacacagtcataatggttatgttaaacaggtc 780
Db 1146 tggctgcatcgtcctgggctcggtctaaatgacacacagtcataatggttatgttaaacaggtc 1205
OY 781 tctggcaggtctgatatgtagtgcacacagaaatcagtcattggtgcatggttggaatgg 840
Db 1206 tctggcaggtctgatatgtagtgcacacagaaatcagtcattggtgcatggttggaatgg 1265
OY 841 caccactcctgaagtgacactcaatatactcctcgaaggtcaacatcttctgtgagaacaa 900
Db 1266 caccactcctgaagtgacactcaatatactcctcgaaggtcaacatcttctgtgagaacaa 1325
OY 901 tggccaggggtctccttggaatctcgcgaataacttcttcttctgctcaaaacttgat 960
Db 1326 tggccaggggtctccttggaatctcgcgaataacttcttcttctgctcaaaacttgat 1385
OY 961 gggccttgagacagcttctactggtttgtcatatctcttcccaacaacatga tggcatgga 1020
Db 1386 gggccttgagacagcttctactggtttgtcatatctcttcccaacaacatga tggcatgga 1445
OY 1021 agcttatgtccaagtagacagctgttccagaggaaccccaactacgaatgaaataatga 1080
Db 1446 agcttatgtccaagtagacagctgttccagaggaaccccaactacgaatgaaataatga 1505
OY 1081 aggaagcggaagactatgatatgatacttactgattcttgaatgtgagtgtgcaagtttga 1140
Db 1506 aggaagcggaagactatgatatgatacttactgattcttgaatgtgagtgtgcaagtttga 1565
OY 1141 tgaatgaacactcctcctcttaaccaaatcgctcagtttgccagaagaagcatctctaaac 1200
Db 1566 tgaatgaacactcctcctcttaaccaaatcgctcagtttgccagaagaagcatctctaaac 1625
OY 1201 ttgggtacatlatcacctgtcgtcgaagagagagactgggactatgctccctagtctctgc 1260
Db 1626 ttgggtacatlatcacctgtcgtcgaagagagagactgggactatgctccctagtctctgc 1685
OY 1261 ccccggttgacagaagtataaagttcaatttgaacaatggtccctcgaaggttggtag 1320
Db 1686 ccccggttgacagaagtataaagttcaatttgaacaatggtccctcgaaggttggtag 1745
OY 1321 gaagtaacaaaagtcggaattatgcatcacacagatgaaacctttaagactctggaagc 1380
Db 1746 gaagtaacaaaagtcggaattatgcatcacacagatgaaacctttaagactctggaagc 1805
OY 1381 tactcagcatgaaatcaggaatcttgggaacttactcttatggggaagttggagacacact 1440
Db 1806 tactcagcatgaaatcaggaatcttgggaacttactcttatggggaagttggagacacact 1865
OY 1441 gtgtattatatttaagaatcaagaacagacatatataacttacccttaaggaaatcac 1500
|||||
Db 1866 gtgtattatatttaagaatcaagaacagacatatataacttacccttaaggaaatcac 1925
OY 1501 tgaatgcgtctccttctgtaattcaagagatttaaccaaaaggtgttaaaacatttgaagatt 1560
Db 1926 tgaatgcgtctccttctgtaattcaagagatttaaccaaaaggtgttaaaacatttgaagatt 1985
OY 1561 tccaattctgcagagagaatattcaaatataaataagacagtgactgttagaagaatgggcc 1620
Db 1986 tccaattctgcagagagaatattcaaatataaataagacagtgactgttagaagaatgggcc 2045
OY 1621 aactaaatcagatctctggtggtcgtgagccgtattactctagtttctgttaattagaag 1680
Db 2046 aactaaatcagatctctggtggtcgtgagccgtattactctagtttctgttaattagaag 2105
OY 1681 agatctagcttcagactcatcttggtccctccctccatcgtgtatcaaaaagatcttaagata 1740
Db 2106 agatctagcttcagactcatcttggtccctccctccatcgtgtatcaaaaagatcttaagata 2165
OY 1741 aagsgaaaaccagataatgltcagaacagaggaatgltcatctcgttctgtatgtatga 1800
Db 2166 aagsgaaaaccagataatgltcagaacagaggaatgltcatctcgttctgtatgtatga 2225
OY 1801 gaaccgaagctggtactcctcaagaagaatatcaacgcttctcccaatccagctggagt 1860
Db 2226 gaaccgaagctggtactcctcaagaagaatatcaacgcttctcccaatccagctggagt 2285
OY 1861 gcaacttgagatccagaagttccaaagcctccaaatcatgacacagcatcaatgtgcatgt 1920
Db 2286 gcaacttgagatccagaagttccaaagcctccaaatcatgacacagcatcaatgtgcatgt 2345
OY 1921 ttttgataagtttgcaatgltcagttgttggatcagltgaggtggtgcatcgtgtatcttaag 1980
Db 2346 ttttgataagtttgcaatgltcagttgttggatcagltgaggtggtgcatcgtgtatcttaag 2405
OY 1981 cattggagacacagactgactcctcttctgctcttcttctcggatataccttcaaaaacaa 2040
Db 2406 cattggagacacagactgactcctcttctgctcttcttctcggatataccttcaaaaacaa 2465
OY 2041 aatgltcatatgaagacacactcaacactatcccatcttccagggagaactgtctctcatgtc 2100
Db 2466 aatgltcatatgaagacacactcaacactatcccatcttccagggagaactgtctctcatgtc 2525
OY 2101 gatggaaaaccaggtctatagattcttgggtgtccaaactcagactcttggaaacagag 2160
Db 2526 gatggaaaaccaggtctatagattcttgggtgtccaaactcagactcttggaaacagag 2585
OY 2161 catbacgccttactggaagtttctagtttgtagcaaaaacacggtgattattagcagga 2220
Db 2586 catbacgccttactggaagtttctagtttgtagcaaaaacacggtgattattagcagga 2645
OY 2221 cagttatgaagatatttcaagcatactgtcgtgagttaaaacaatgcatcttgaaaccaagaag 2280
Db 2646 cagttatgaagatatttcaagcatactgtcgtgagttaaaacaatgcatcttgaaaccaagaag 2705
OY 2281 ctctccccaagaattcaagaacacctagcactagcgaagaatthaatgccaacccac 2340
Db 2706 ctctccccaagaattcaagaacacctagcactagcgaagaatthaatgccaacccac 2709
OY 2341 agtcttgaaagccatcaacggaatcaactcgtactactcttcagttcagttcagaaga 2400
Db 2710 agtcttgaaagccatcaacggaatcaactcgtactactcttcagttcagttcagaaga 2747
OY 2401 aattgcatatgatacatcatcatcagttgaaatgaaagaaggaatttgacattatga 2460
Db 2748 aattgcatatgatacatcatcatcagttgaaatgaaagaaggaatttgacattatga 2807
OY 2461 tgaagatgaaatcaagagcccgagcttccaaaagaacacagacataatttatgtc 2520
Db 2808 tgaagatgaaatcaagagcccgagcttccaaaagaacacagacataatttatgtc 2867
OY 2521 tgcagtgagaggtcttgggataatggagatgagtagtctcccaatggttcttaagaacag 2580
Db 2868 tgcagtgagaggtcttgggataatggagatgagtagtctcccaatggttcttaagaacag 2927
```

QY 2581 ggtcagagtgagcagtgccctcagttcaagaaagtgttttccagaattactatg 2640  
|||||  
Db 2928 ggtcagagtgagcagtgccctcagttcaagaaagtgttttccagaattactatg 2887  
QY 2641 cctccttactcagccttataccgtggaagaaactaaatgaacatttggactcctgggccc 2700  
|||||  
Db 2988 cctccttactcagccttataccgtggaagaaactaaatgaacatttggactcctgggccc 3047  
QY 2701 atatatagaagcagaagttggaagataatcatgtaactttcaagaatcagcctctcg 2760  
|||||  
Db 3048 atatatagaagcagaagttggaagataatcatgtaactttcaagaatcagcctctcg 3107  
QY 2761 tccctatccttctatctcagccttatttcttctgaaggaaatcagaagcagaagcaga 2820  
|||||  
Db 3108 tccctatccttctatctcagccttatttcttctgaaggaaatcagaagcagaagcaga 3167  
QY 2821 acctagaanaaacttctcagccttaatagaacaaacttacttcttgaagtgaca 2880  
|||||  
Db 3168 acctagaanaaacttctcagccttaatagaacaaacttacttcttgaagtgaca 3227  
QY 2881 tcatatggcccccactaaagatgagtttgactgcaagccttggtctatttctctgagt 2940  
|||||  
Db 3228 tcatatggcccccactaaagatgagtttgactgcaagccttggtctatttctctgagt 3287  
QY 2941 tgaacctggaanaagatgctgacccagcctgagcttgagccctctgctgcacacaa 3000  
|||||  
Db 3288 tgaacctggaanaagatgctgacccagcctgagcttgagccctctgctgcacacaa 3347  
QY 3001 cacactgaacctgctcatalggaagacaaatgacagtaacagaaatttgcctgttttccac 3060  
|||||  
Db 3348 cacactgaacctgctcatalggaagacaaatgacagtaacagaaatttgcctgttttccac 3407  
QY 3061 catcttctgtagagcacaagaagctggtacttccatgtaaatatggaagaagcagcagc 3120  
|||||  
Db 3408 catcttctgtagagcacaagaagctggtacttccatgtaaatatggaagaagcagcagc 3467  
QY 3121 tccctcgaatccagatggaagatcccaactttaaagaagatacgcctccatcagat 3180  
|||||  
Db 3468 tccctcgaatccagatggaagatcccaactttaaagaagatacgcctccatcagat 3527  
QY 3181 caatggtctacataatgatactactgctgtagtaatggtcagagatcaaaagattcg 3240  
|||||  
Db 3528 caatggtctacataatgatactactgctgtagtaatggtcagagatcaaaagattcg 3587  
QY 3241 atggtatctgctcagcagatggtgagcaatgaaacatccatcttcatcttcaagtga 3300  
|||||  
Db 3588 atggtatctgctcagcagatggtgagcaatgaaacatccatcttcatcttcaagtga 3647  
QY 3301 tgtgttcactgtaagaaaaaagagagataaataatgagcctgtacaaatctctacag 3360  
|||||  
Db 3648 tgtgttcactgtaagaaaaaagagagataaataatgagcctgtacaaatctctacag 3707  
QY 3361 tgtttttgagacagtggaatgtttacatccaaagcttgaatttggcgggtggaatgct 3420  
|||||  
Db 3708 tgtttttgagacagtggaatgtttacatccaaagcttgaatttggcgggtggaatgct 3767  
QY 3421 tatgtgagcagatctacatgctggtggaatgagcaacttttttgggtgtacagcaataatg 3480  
|||||  
Db 3768 tatgtgagcagatctacatgctggtggaatgagcaacttttttgggtgtacagcaataatg 3827  
QY 3481 tcaagatcccttgggaatggtctctgagacatctagaattttcaatatcagactcag 3540  
|||||  
Db 3828 tcaagatcccttgggaatggtctctgagacatctagaattttcaatatcagactcag 3887  
QY 3541 acaaatgagacagtggtgccccaaagctggtccagactctatctccggtcaatcaatgc 3600  
|||||  
Db 3888 acaaatgagacagtggtgccccaaagctggtccagactctatctccggtcaatcaatgc 3947  
QY 3601 ctggaagcacaagaagccttttcttctgtagcaagtgatcgtgtgacccaatgattat 3660  
|||||  
Db 3948 ctggaagcacaagaagccttttcttctgtagcaagtgatcgtgtgacccaatgattat 4007

QY 3661 tcaagcatalcaagaccaggtgtcccgctcagaagttctccagcctctacatctcagtt 3720  
|||||  
Db 4008 tcaagcatalcaagaccaggtgtcccgctcagaagttctccagcctctacatctcagtt 4067  
QY 3721 tatcatcatgtatagctcttgaatggaagaatggtgcaacttctcagaagaattccactg 3780  
|||||  
Db 4068 tatcatcatgtatagctcttgaatggaagaatggtgcaacttctcagaagaattccactg 4127  
QY 3781 aaacttaatggtctcttcttggcgaatgtggaattcatcttggaataaacacataatttta 3840  
|||||  
Db 4128 aaacttaatggtctcttcttggcgaatgtggaattcatcttggaataaacacataatttta 4187  
QY 3841 cctccaatattgtctcgatataccggtttgacccaactcatatagacttgagcagca 3900  
|||||  
Db 4188 cctccaatattgtctcgatataccggtttgacccaactcatatagacttgagcagca 4247  
QY 3901 tcttcagatgagttgagtgtggtctgtatataatgltgcaagatgcatctgggaattgga 3960  
|||||  
Db 4248 tcttcagatgagttgagtgtggtctgtatataatgltgcaagatgcatctgggaattgga 4307  
QY 3961 gagtaagaacatactcagatgacagattactgtcttacttacttaccataatgtttgc 4020  
|||||  
Db 4308 gagtaagaacatactcagatgacagattactgtcttacttacttaccataatgtttgc 4367  
QY 4021 caactgtctccttcaaaagctcgacttcaactccaaggagagatgaatgctggagacc 4080  
|||||  
Db 4368 caactgtctccttcaaaagctcgacttcaactccaaggagagatgaatgctggagacc 4427  
QY 4081 tcaagtgaaatccaaagaatggtgctgcaagtgaacttccagaagaacataatgaatcac 4140  
|||||  
Db 4428 tcaagtgaaatccaaagaatggtgctgcaagtgaacttccagaagaacataatgaatcac 4487  
QY 4141 aggaagtaactactcagaggtgaaatctcgtctaccagagatgatagtgaagagttcct 4200  
|||||  
Db 4488 aggaagtaactactcagaggtgaaatctcgtctaccagagatgatagtgaagagttcct 4547  
QY 4201 catctccagcagtcagatgagccatcagtgagctctcttttccagaatggaagtaaa 4260  
|||||  
Db 4548 catctccagcagtcagatgagccatcagtgagctctcttttccagaatggaagtaaa 4607  
QY 4261 gtttttcagggaatacgaactccttccacactgtgtgtaactctcttagaccacagtt 4320  
|||||  
Db 4608 gtttttcagggaatacgaactccttccacactgtgtgtaactctcttagaccacagtt 4667  
QY 4321 actgactcgtactcctcgaaattcaaccocagagttggtgacacagatggtcctgaagat 4380  
|||||  
Db 4668 actgactcgtactcctcgaaattcaaccocagagttggtgacacagatggtcctgaagat 4727  
QY 4381 ggaagttctggggtgagcagacagacacttactga 4417  
|||||  
Db 4728 ggaagttctggggtgagcagacagacacttactga 4764

RESULT 8  
US-08-882-083-1  
: Sequence 1, Application US/08882083  
: Patent No. 5869292  
: GENERAL INFORMATION:  
: APPLICANT: VOORBERG, Johannes J.  
: TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY  
: NUMBER OF SEQUENCES: 17  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Foley & Lardner  
: STREET: 3000 K Street, N.W., Suite 500  
: CITY: Washington  
: STATE: D.C.  
: COUNTRY: USA  
: ZIP: 20007-5109  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.30

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/882,083
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/558,107
FILING DATE: 13-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: ISACSON, John P.
REGISTRATION NUMBER: 33,715
REFERENCE/DOCKET NUMBER: 30472/212
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEO ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5035 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 35..5017
US-08-882-083-1

Query Match      84.8%  Score 3743.8:  DB 2:  Length 5035:
Best Local Similarity 87.5%:  Pred. No. 0:
Matches 4355:  Conservative 0:  Mismatches 52:  Indels 570:  Gaps 2:

QY 1 catgcaaatagaagctctccacctgctctctctctgctgctctcttgcatctgctctatgctc 60
DB 34 CATGGAATAGAGCTCTCCACCTGCTTCTTCTGCTGCTTGGCATTCCTGTTAGTGC 93
QY 61 caccagaagatactactcctggtgctgagtggaactgcatggtgactatatagcaagtgatct 120
DB 94 CACCAAGAATAGTACTACTCTGGGTGAGTGAATCTCATGGGACATATGCAAAATGATCT 153
QY 121 cggtagagctcctgctggaagcaagatctctcctcctagagtgccaacatcttccatcaa 180
DB 154 CGGTGAGCTGCTCTGTGAGCCAGATTTCTCTCTAGAGTGCAGAAATCTTTCCATTCAA 213
QY 181 caccctagctgctgtaacaaaagaactcgtttgtagaatltcaaggttcaacctttcaaat 240
DB 214 CACCTCAGTCTGTACAAAAGACTCTGTTGTGAAATTCACGGATCACCTTTCAACAT 273
QY 241 cgttaagccaagccaccctggtgagtgctgctagtgctcctaccatccaagctgagttta 300
DB 274 CGCTAAGCCCAAGGCCACCTGGATGGGTCTGTAAGTCTTACCAATCCAGGCTGAGGTTTA 333
QY 301 tgaataagtgatcatltaacttaagaacatlggcttccatccctgctgaagcttcaagctctgt 360
DB 334 TGATACAGTGTGATTTACATTTAAGAACATGCTTCCCATCTCTCAGTCTTCTGCTGT 393
QY 361 tggtagtactctactggaagctctctggaagagctgtaaatagatgatacagaacagtcgaag 420
DB 394 TGGTGTATCTACTACTGGAACCTTCTGAGGAGTGAATATGATGATCAGACCACTCAAG 453
QY 421 ggaagaagaagatgataaagctctcctcctggtggaagccatacatatgcttggcaagctc 480
DB 454 GGAAGAAGAAGATGATTAAGTCTTCCCTGCTGGAAGCCATATGCTGGCAGGCTCT 513
QY 481 gaaagaagaatggtccaaatggcctctgagccactgtgcttaaccactcatacttcttcca 540
DB 514 GAAAGAAGAATGCTCAATGGCTCTTGACCACTGTGCTTACCTACATATCTTTCTTCA 573
QY 541 tgtgagcctgttaaaagacttgaatcaggacctcatctggagccctactagtaigtataga 600
DB 574 TGTGAGACTGTGTAAGACTTGAATTCAGGCTCATTTGAGCCCTACTAGATATGAGGA 633
QY 601 aggagactgagcagaagaaagacacagaccttgacacaaattatactacttcttgctgt 660
|||||

DB 634 AGGAGTCTGCGCCAAGAAAAAGACACAGACCTTGCCACAAATTTATACCTTTTGTCTGT 693
QY 661 attgatgaagggaaaagtgtgacactcagaacaaagaactccttgatgagatagga 720
DB 694 ATTGATGAAGGAAAGTTGGTGGCACTCAGAAACAAACAACTCTGTGATGAGATAGGA 753
QY 721 tgtgtcatctgctcggcctggcctaaatgacacagtcgaatggttatgttaaagctc 780
DB 754 TGTGCAATCTGCTGGGCTGGCTAAATGACACAGTCAATGTTATTAACAGGCTC 813
QY 781 tctcagaggtctgattgtagcagaagaatgaatgctatctggcatgtatggaatgg 840
DB 814 TCTGCCAGGTCTGATTTGATGCCACAGGAATCACTATTTGGCATGTGATTTGGAATGG 873
QY 841 caccactctgaagtgacactcaatatctccctcgaaggtcacacacttcttctgagaaca 900
DB 874 CACCACTCTGTAAGTGCACCTCAATATTCCTGGAAGGTCAACATTTCTTGTGAGAACCA 933
QY 901 tcgcaagcgctccttggaaatctgcacaatacttccctactgctcacaaccttggat 960
DB 934 TCGCAGGCGTCTTGGAAATCTGCCAATTAATTCCTTACTGCTCAACACTTGTAT 993
QY 961 ggaacctggagaagtttctactggtttgtcaatctcttccaccacaatgagtcatga 1020
DB 994 GGACCTTGAGACATTTCTACTGTTTGTGATATCTCTTCCACCAACATGATGCAATGA 1053
QY 1021 agctatgcaaaatgacagactgtccagaggaaccccaactcgaatgaaataatga 1080
DB 1054 AGCTTATGTCAAGTACACAGCTGTCCAGAGGAACCCCACTCAGATGAAATATATGA 1113
QY 1081 aagaagcggaagactatgatagtacttactgatatctgaaatgagatgtgcaagtttga 1140
DB 1114 AGAAGCGGAAGACTATGATGATGATCTTACTGATTCGAATGAGATGCTGACGTTTGA 1173
QY 1141 tgaatgaacactctcctccttccatccaaatctgctgaatgccaagaagcatcctaagc 1200
DB 1174 TGATGACAACTCTCTCTTCTTATCCAAATTCGCTGATGTCAGAGAGATCTTAAAC 1233
QY 1201 ttggtagatcatcatctgctcgtgaagagagagactgggaactgctccttaagctcgc 1260
DB 1234 TTGGGTACATTACATTTCTGCTGTGAAGAGAGACTGGGACTATGCTCTTACTCTCCG 1293
QY 1261 ccccgatgacagaagtataaagaatcaatatltgaacaatgcccctcagcgatgtgag 1320
DB 1294 CCCGATGACAGAAATTATAAAGTCAATTTTGAACAAAGGCCCTCAGGGGATTTGTTAG 1353
QY 1321 gaagtacaaaagaatccgattatgcatatacagaatgaacactttaaagctctgaagc 1380
DB 1354 GAAGTACAAAAGATCCGATTTATGSCATACAGATGAACCTTTAAGACTGTGAAGC 1413
QY 1381 tatcagcatgaaatcaaggaatcttggagcttacttatagggaagtltggaacacact 1440
DB 1414 TATTCAGCATTAATCAGGAATCTTGGGACCTTTACTTTATGGGAGTTGAGACACACT 1473
QY 1441 gttgattatalttaagaatcaagaacagaacatataacatccctcaagaaatcac 1500
DB 1474 GTTGATTATATTTAAGATCAAGCAACAGCATATATACATCTACCCACAGGAATCAC 1533
QY 1501 tgaatgctcgtccttctgattccaaggaatatacacaagaaggttaaacattltgaagatt 1560
DB 1534 TGATGTCCGCTCTTGTATTCAAAGGAATTAACAAAAGGTGTAAACATTTGAAGGATTT 1593
QY 1561 tccaattctgccaaggaagaatatccaataataatggaagtgactgtagaagaatggcc 1620
DB 1594 TCCAAATCTGCCAGGAAATATTTCAATATTAATGACAGTCACTGTGAACATGGGCC 1653
QY 1621 aactaaatcagatcctggtgctgacccgtatacttactagtttcgttaataatggagag 1680
DB 1654 AACTAAATCAGATCTCGTGGTCCGTGACCCGCTATTTACTAGTTTGGTTAATATGAGAG 1713
QY 1681 agatcctagctcaggaactcaatggcctctcctcactatctgtataaagaatctgtataca 1740
DB 1714 AGATCTAGCTTCAGGACTCATTTGGCCCTCTTCATCTCATCTGTACAAAGAAATCTGTATGA 1773
|||||
```



1741 aagaggaaacgaataatgtcaagacaagagaaatgcatcctgttttctgtattga 1800  
1774 AAGAGAAACCGAATATGTCAACAGAGAAATGTCATCTGTTTCTGTATTATGA 1833  
1801 gaaccgaactggttacctacagagaataatacaacgcttctccccaatcagctgaat 1860  
1834 GAACCGAAGCTGTACTCTCAAGAGAAATATCAACGCTTCTCCCAATCCAGCTGAGT 1893  
1861 gcaagcttggagatccagagatcccaagcctccaaatcatgacagacaataatggtatgt 1920  
1894 GCAAGCTTGGAGATCCAGAGATCCAAACCTCCAAACATCATGACAGCATCAATGGCTATGT 1953  
1921 ttcttgatgttgcagttctcagttgttcttgatggatggagactgttatacttaag 1980  
1954 TTTTGATGTTTGCACTTGTCTCACTTTGTTGATGAGGTGGCACTGTGATCTTAAAG 2013  
1981 catlgtgagacagactgacttctctgtctctctctctgatalacttcaaacaca 2040  
2014 CATTGAGACACAGACTGACTTCTCTTCTCTCTCTCTGATATACCTTCAACACAA 2073  
2041 aatgtctatgaagacaactcaacctatcccatctcaaggagaaactgtcttcaatgc 2100  
2074 AATGCTATGAAGACACACTCACTATTCCTATCTCAGAGAAAGTCTTATGTC 2133  
2101 gatgagaaacccaagtctatlgatltctgggtgcccacaactcagacttccggaacagag 2160  
2134 GATGAGAAACCCAGGCTATGAGATTCCTGGGTGCCAACACTCAGACTTTCGGAACGAGG 2193  
2161 catgaccgcttactgaaggttctcagttgttgaacaagaactggtgatattacaagga 2220  
2194 CATGACCGCTTACTGAAGGTTTCTAGTGTATTCAGAGGGGGAGAGAGACACACTA 2253  
2221 cagttatgaagatattca-----gcatactgctgaatgaacaacatgc 2265  
2254 TCTGGACCTGGAGAGATATTCAGTGAAGACAGACACTGATGACATGCTGCACAGTCT 2313  
2266 catlgaacaaagaactctccagaatccagaacacccctagcactgaagcaaaagaact 2325  
2314 GATTGAACCAAGAGCTTCTCCAGATTCAGACACCTAGCACTAAGCAAAAGCAATTT 2373  
2326 taatgcac----- 2334  
2374 TATGCGACACCAATTCAGAAATATGACATAGAGAACTGACCTTGGTTGSCACAG 2433  
2335 ----- 2334  
2434 AACACCTATGCTAAATATCAAAATGTCTCTAGTATTTGTTGATGTCTCTGGGACA 2493  
2335 ----- 2334  
2494 GAGTCTACTCCACATGGGCTATCTTATCTGATCTCCAGAAAGCCAAATATGAGACTTT 2553  
2335 ----- 2334  
2554 TTCTGATGATCATCATGCTGAGCAATATACAGTATATACAGCCTGTCTGAAATGACACA 2613  
2335 ----- 2334  
2614 CTTACAGGCACAGCTTCATCAGAGTGGGACATGATTTACCCCTGAGTCAGGCTCCA 2673  
2335 ----- 2334  
2674 ATTAAAGATTAAATGAGAAACTGGGACAACCTGAGATCTCTTGGTGGATTAACCACTA 2733  
2335 ----- 2334  
2734 TGGTACTCAGATACCAAAAGAGATGGAATCCCAAGAGACTCACAGAAAAAAGACAGC 2793  
2335 ----- 2334  
2794 TTTTAAGAAAAAGATACCATTTTGTCTCTGAAAGCTTGTGAAGCAATCATGCAATAGC 2853

2335 ----- 2334  
2854 AGCAATTAATGAGGACAAATAAGCCCGAAATAGAGTACCTGGGCAAAAGCAAGTAG 2913  
2335 ----- 2334  
2914 GACTGAAGGCTGTGCTTCACAAACCCAGAGCTTGAAAGCGCATCAACGGGAAATTAAC 2973  
2371 tggactactcttccagtcagatcaagaagaaattgactatgataacatacaatga 2430  
2974 TCGTACTACTCTTCAGTCACTGATCAAGAGAAATTTGACTATGATGAATACCATACACTTGA 3033  
2431 aatgaagaagaagatttgcattatgaatgaatgaatgaatgaatgaatgaatgaatgaat 2490  
3034 AATGAAGAAAGAAATTTTGCATTTATGATGAGATGAAATTCAGAGCCCGGACGCTT 3093  
2491 tcaaaagaaacacagacactatttlatlgtcgaatgagagagctctgagatlatggat 2550  
3094 TCAAAGAAACACAGACACTATTTATGCTGCAGTGAAGAGGCTCTGGGATATGAGAT 3153  
2551 gagtagctcccaatgcttctcaagaacagagctcagatgagatgctccatgaatcaa 2610  
3154 GAGTACTCCCGCACATGTTCTTAAGAAACAGGGCTCAGAGTGGCAGTGTCTCCATGTTCAA 3213  
2611 gaaagtgttttccagaatctatgaatgctccttactcagcccttatacsgtggaga 2670  
3214 GAAAGTTGTTTCCAGGAATTTACTGATGAGCTCTTACTCAGCCCTTATACCGTGAGGA 3273  
2671 actaaatgaacatttggagctcctggggccatataaagaacagaatltgaagataat 2730  
3274 ACTAATGAACATTTGGAGCTCCTGGGCCATATATTAAGCAGAAAGTTGAATATAT 3333  
2731 catgtaacttccaagaatcagagctctcgtccctatcccttactatcagcttattc 2790  
3334 CATGCTAATTTTCAGAAATCAGGCTCTGCTCCCTATTCCTTATTTAGCCTTATTTTC 3393  
2791 tlatgagaagaatcagagcaagagcagaacccaaagaacttgcagaagcaatga 2850  
3394 TTATGAGGAAGATCAGAGGAGAGAGACAGACTAAGAAATCTTGTCAGGCTATATGA 3453  
2851 aaccaaacttacttcttgaagaatgtcaacatcatatgagacccaactaaagatgaattga 2910  
3454 AACCAAACTACTTCTTGGAAAGTGCACATCATATGAGCCACCACTAAGATGAGTTTGA 3513  
2911 ctgcaagcctgggcttattctctgattgttgaactggaagaagatgtgaactcagagct 2970  
3514 CTGCAAAAGCTGGCTTATTTCTGATGTGACCTGGAAGAAAGATGTGCACAGGCTCT 3573  
2971 gattgagccctctgtgtcgcacactaaacagactgaacccctgctcatgaggagagaagt 3030  
3574 GATTGACCCCTTCTGTGCTGCCACACTTAACACACTGAACCTGCTCATGGAGACAAGT 3633  
3031 gacagtlacagaatltgctctgttcttccacactlttgaatgaagcaaaagctgactt 3090  
3634 GACAGTACAGGAATTTGCTCTGTTTCAACCATCTTGAATGAACCAAAAGCTGTACTT 3693  
3091 cactgaataatgaagaagaactcgaaggtccctcgaatataccagaatgaatccac 3150  
3694 CACTGAAATATGGAAGAAAGAACTGCAGGGCTCCCTCAATATCCAGATGGAAGATCCAC 3753  
3151 tttaaagaatlatcgttccatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 3210  
3754 TTTTAAAGAAATTAATGCTCTCATCAATCAATGAGCTACATATATGATACCTACTGAG 3813  
3211 cttagtaatgagctcgaagacaaaggaatcgaatgatactcgtcgaatgagagcaatga 3270  
3814 CTTAGTAATGGGCTCAGAGATCAAAAGATTTGATGATGATGATGATGATGATGATGATG 3873  
3271 aaacatcaatctatcatcttcaatgagatgagatgagatgagatgagatgagatgagatg 3330  
3874 AAACATCAATTTCTATTTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 3393  
3331 taaaatgagacgttacaatctctatccaggtgttlttgaagaagtgaatgttaccatc 3390



```
|||||
Db 3934 TAAATGGCACTGTACATCTATCCAGGTTTTGAGACATGGAATGTACATC 3993
QY 3391 caaagctgaatttgcgggttgaaatgcttaattgagcaatctacatgcttgatgag 3450
Db 3994 CAAGAGTGGAAATTTGGGGGTGGAAATGCTTATTGGGAGCATCTACATGCTGGAGTAG 4033
QY 3451 cacacttttcttggtgtacagcaataagtgtaagactcccttggaatggcttctgaca 3510
Db 4054 CACACTTTTCTGCTGTAGACGAATAGTGACAGATCCCTGGGAATGGCTTCTGGACA 4113
QY 3511 catagaagatttcagattcagcttcaggaacaatagagatgggccccaaagctggc 3570
Db 4114 CATTAGAGATTTTCAGATTACACTTCAGAGACATATGAGACAGAGGCCCAAGCTGGC 4173
QY 3571 cagacttcatatttcgagatcaatcaatgcttgagacccaagagaccccttcttgat 3630
Db 4174 CAGACTTCATTATTCCGGATCAATCAATGCTGAGACCAAGAGGCCCTTTCTTGAT 4233
QY 3631 caagtgatctgttgacacaaatgattatcaagcatcaagaaccaggggtgcccgtca 3690
Db 4234 CAAGGTGATCTGTTGGCACCAATGATTAATTCAGGCATCAAGACCAGGCTGCCGTCA 4293
QY 3691 gaagttctcagcctctacatctctcagttatcatatgatatagttcttgatggaa 3750
Db 4294 GAAGTCTCCAGCCTCTACATCTCTCAGTTATCATCATGTATAGTCTTGATGGAGAA 4353
QY 3751 gtcgacagacttcaggaatccacttcagaaacttaagttctctcttgcaatgtgga 3810
Db 4354 GTGGCAGACTTATCGAGAAATTCACATGACCTTAATGTCCTCTTGGCAATGTGGA 4413
QY 3811 ttcatctgggaataaaacaatatltttaaaccctcaatatgtctcgatacatcgtt 3870
Db 4414 TTCATCTGGGATTAACACAAATATTTTAAACCCCTCAATATTGCTGATACATCCCTTT 4473
QY 3871 gaaccaactcatatagacttgccagcactcttcgcatgagttgagtgctgatt 3930
Db 4474 GACCCCAACTCATTAATAGCAATTCGCACACTCTTCGATGAGTTGAGGCTGTGATTT 4533
QY 3931 aataatctgaagcatgcacttggaatgagagatgaagcaatcaatgagacagattac 3990
Db 4534 AATATGTCGAGATGCGCATTTGGGAATGAGAGATGAAGCAATATCAGATGACAGATTAC 4593
QY 3991 tgcttaccctactttaaacaatagtlttgcaacctgtctcctcaaaagctgacttca 4050
Db 4594 TGCTTACTCTTACTTAACCAATATGTTTGCACCTGCTCTCTCAAAAGCTCGACTTCA 4653
QY 4051 cctccaaaggagagatgaatgctgagaaactcaggtgaataatccaaaagatggctgca 4110
Db 4654 CCTCCAAAGGAGAGATGATGCTGAGACCTCAGTGAAATTAATCAAAAGATGGCTGCA 4713
QY 4111 agtgaacttccaaagaacaatgaagaatcagagagtaactcagagagaaatctct 4170
Db 4714 AGTGAATTCGAAGAACAATGAAGTACAGAGATTAATCAAGGAGTAATCTCT 4773
QY 4171 gcttaccagaatgatatgtgaaggagttctcatctcagcagatcaagaatggcactg 4230
Db 4774 GCTTACCACATGTATGTGAAGAGTCTCATCTCCAGCAGTCAAAATGTGCTCATGTG 4833
QY 4231 gactcctcttttcagaatgcaaaagtaaggttttccaaggaaatcaagaactctcac 4290
Db 4834 GACTCTCTTTTTCAGATGCAAAAGTAAGGTTTTTCAGGGAATCAAGCTCTCTTAC 4893
QY 4291 acctgtgtaactctcttagaccacacgcttactgactcgtactcgaatltcaaccaca 4350
Db 4894 ACCTGTGTAATCTCTTAACACCAACCGTACTGCTGCTGCTTGAATTAACCCCA 4953
QY 4351 gagttggtgacacagattgacctgagatgaggttctggtctcgagagacagaact 4410
Db 4954 GAGTTGGTGCACACAGATTGCTCTGAGAGATGAGGTTCTGGGCTGCAGGACAGACCT 5013
QY 4411 ctactga 4417
|||||
```

```
Db 5014 CTACTGA 5020

RESULT 9
US-08-558-107-1
; Sequence 1, Application US/08558107
; Patent No. 5910481
; GENERAL INFORMATION:
; APPLICANT: VOORBERG, Johannes J.
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; City: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/558,107
; FILING DATE: 13-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: ISACSON, John P.
; REGISTRATION NUMBER: 33,715
; REFERENCE/DOCKET NUMBER: 30472/212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SRO ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5035 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 35..5017
; US-08-558-107-1

Query Match 84.8%; Score 3743.8; DB 2; Length 5035;
Best Local Similarity 87.3%; Pred. No. 0;
Matches 4365; Conservative 0; Mismatches 52; Indels 570; Gaps 2;

QY 1 catgcaatagagcttccacacgtcttcttctgtgcttcttgagattcgtcttagtgc 60
Db 34 CATGAATAGAGCTCTCCACACTCTCTTCTGTGTGCTTTTGGATTCGCTTAACTGTC 93
QY 61 caccagaagatactactggtgtgagtgagtggaactgtcatggaactatataatgaatgatct 120
Db 94 CACCAAGAATACACTGCTGATGAGTGAAGTCAATGAGACTATATGCAAGATGATCT 153
QY 121 cgtgagactgctgtgagcgaagatttctctctctagatgagtgcaaaatcttctccatca 180
Db 154 CGGTGACTGCTGTGGACCAAGATTCTCTCTAGAGTCCAAATCTTTTCCATTCA 213
QY 181 caactcgtgtgtacaaaagaactcgttgtgagaattcaagcttcaacttcaaat 240
Db 214 CACTCTAGTGTGTACAAAAGATCTGTTGTAGATTTCAGGATTCACCTTTTCAACAT 273
QY 241 cgttaagccaagggcacccctgagatgggtctgctagtgctactaccacacaggtgagttta 300
Db 274 CGTAAAGCAAGGCAACCTGAGATGGGTCTGAGTCTTACCATCAAGGCTGAGGTTTA 333
QY 301 tgataagtggttatactacttaagaacatggttcccatccctggtgaagtttatagtctgt 360
|||||
```

Db 334 TGATACAGTGTCATTAAGAACATGCGTTCCCATCTGTGACGTCTGATGCTGT 393  
QY 361 tgggtatctactaggaaagcttctgaggagctgaatgatgatagaagcagtcgaag 420  
Db 394 TGGGTATCTCTACTGGAAAGCTTGTAGGAGCTGGAATGATGATAGACACGTCAAG 453  
QY 421 ggaagaaagatgatataagctctccctggctggaagcatatcatatctcggagctcct 480  
Db 454 GGAGAAAGAAATGATATAAGTCTTCCCTGGTGGAGGCATACATATCTGCGCAGGTCT 513  
QY 481 gaaagaaatggtccaaatggcctctgagccacatgctcctaactatcatcttctca 540  
Db 514 GAAAGAAATGTCCTCAATGTCCTGTGACCACCTGTGCTTACCTACATATCTTCTCA 573  
QY 541 tgggagactggtataagacttgaaatcccatcttgaggccctactaatgatagaaga 600  
Db 574 TGTGGACCTGTGTAATAACATTGAATTCAGGCTCATTTGGAGCCCTACTATGATGAAGA 633  
QY 601 agggagctgagccaaagaaagacacagaccttgacaaatctataactcttctgtgt 660  
Db 634 AGGGAGCTGTGGCAAGAAAGACACAGACCTTGACAAATTTATACTACTTTTGTCTGT 693  
QY 661 attgatgaagggaagaaagtgtgacactgaagaaagaaactccttgaatgagagatagga 720  
Db 694 ATTGATGAAGGGAAGAAAGTTGGCAGCTCAGAAACAAAGAACTCTGTGATGAGAGTAGGA 753  
QY 721 tggctgatctgctcgggcccggcctaaatgacacacagcttaagtgtaatgaaaggtc 780  
Db 754 TGTGTGATGCTGTGCTGGGCTGGGCTTAAATGACACAGTCAATGTTATGTTAAACAGTTC 813  
QY 781 tctgcaagctgctgatgatgacacagaaatcaagctctatctgacatgtaatgaaatgga 840  
Db 814 TCTGCCAGGCTGTGATTTGATGACCAAGAAATAGTCTATTTGGCATGTGATTTGGATGGG 873  
QY 841 caccactcctgaagtgacactcaatctcctcgaagylcacacatcttctgtgaggaacca 900  
Db 874 CACCACCTCTGGAAGTGCACATATTTCCCTGGAAGGTCAACATCTTCTTGTGAGGAACCA 933  
QY 901 tgcgcagaggtcctctggaatctgcgaataacttctactgcttcaaacactcttgat 960  
Db 934 TCGCCAGGCGCTCTGGAAATCTCGCCAAATACCTTTCCTTACGTCAACACCTTTGAT 993  
QY 961 ggaacctgagcaatcttactatgcttctgcatatctctctccacaacatgatgacatgga 1020  
Db 994 GGACCTTGGACATTTCTACTGTTTGTCTATCTCTTCCACCAACATGATGGCATGGA 1053  
QY 1021 agctctatgtaaaagtgaagcagctgtccagaagaaaccccaactaagaaatgaataatga 1080  
Db 1054 AGCTTATGTCAAAAGTACAGAGCTGTCCAGAGAACCCCAACTAGCAATGAATAATATGA 1113  
QY 1081 agaaagagaaagactatgatgatcttactgattctcgaatgatgagtgagtgagttga 1140  
Db 1114 AGAAGCGGAAGACTATGATGATGATCTTACTGATTTCTGAATATGATGTGTCAGGTTTGA 1173  
QY 1141 tgaatgaacatctccttcccttcttcaaatctcagctgacagaagaagatcctaaac 1200  
Db 1174 TGAATGACAACTCTCTCTTCTTATCCAAATTCCTCTCAGTTGCCAAGAGATCTCTAAAC 1233  
QY 1201 ctgggtacatatacatctgctgtcgaagaagagagactggagactatgctccttagctcctgc 1260  
Db 1234 TTGGGTACATATTGATGCTGTCTGAGAGGAGACTGGAGATGATGCTCCCTTACTGCTCGC 1293  
QY 1261 ccccgatgaagaagttataaagttcaataattgaaacaatggccctcagaaggttgtag 1320  
Db 1294 CCCCAGTACAGAAAGTATATAAGTCAATATTTGAACAAATGGCCCTCAGGGAGTTGGTAG 1353  
QY 1321 gaagtaaaaaaagtcogatttlatgycatcacagatgaaaccttgaagctgtgaaagc 1380  
Db 1354 GAAGTACAAAAAGTCOGATTTATGGCATACAGATGAACCTTTTAAAGCTGTGAAGC 1413  
QY 1381 tattcagcatgaaatcaggaactcttggaaccttacttlatgggaaggttggagacacact 1440  
Db 1414 TATTCAAGCATGAATGAGAAATCTTTGGACCTTACTTTATGGGGAAGTTTGAACACACT 1473

QY 1441 gtgatattatlttaagaatcaagcagacccatataaaccttaccctcaaggaatcac 1500  
Db 1474 GTTGATTTATTTATAGATACAGACAGACCATATTAACATCTACCTCAGGAATCAC 1533  
QY 1501 tgaatgctccctcttgatcacaagagattaccaaaggtglaaacaatttgaagattc 1560  
Db 1534 TGATGTCCGTCTTGTATTCAGAGAGATTACCAAAAAGGTAAACATTGGAAGGATTT 1593  
QY 1561 tccaatctgcccagagaaatlatcaaatataaaltgagacatgacatgtagaagatggcc 1620  
Db 1594 TCCAAATCTCCAGAGAAATATTCAAATTAATGAGACAGTACGTGAAAGATGGGCC 1653  
QY 1621 aaactaaatgaatcctcgtggccctgagccgcttactcagtttctgtaatatgagag 1680  
Db 1654 AACTAAATCAGATCTCTGGTGGCTGAGCCCTATTTACTGTATGTTTCTGTATATGAGAG 1713  
QY 1681 agatctagctcagagacttcatgcccctccatcatctgtacaaagatctgtagatca 1740  
Db 1714 AGATCTAGCTTTCAGAGACTCATTTGGCCCTCTCTCATCTGCTACAAAGATCTGTAGATCA 1773  
QY 1741 aagaagaaacagaaataatgacagaagaagaaatgcatccctgttctglatlttga 1800  
Db 1774 AAGAGAAACCAAGATATGTCAAGACAGAGAAATGTCATCTGTGTTTCTGTATTTGATGA 1833  
QY 1801 gaaacgaagctgtgactcctcagagaatatcaaacgttctcccaatccagctgagct 1860  
Db 1834 GAACCGAAGCTGTACTCTCAGAGAAATATCAACGCTTCTCCCAATCCAGCTGGAGT 1893  
QY 1861 gcaagctgagatcccaagatgccaaagctccaaacatcatgcaagcaatcaatgctatgct 1920  
Db 1894 GCAGCTTGAAGATTCAGAGATTCAGAGCTCCACATCATGACACAGCTCATGTGATGT 1953  
QY 1921 tcttgatagttgcaggtgtcaggttctgttgatgaaatgagatgacatgtaactgaag 1980  
Db 1954 TTTTGAATAGTTTGCAAGTTGTGAGTTGTTGATGAGGTGCAATGCTGATCTTCTTAAG 2013  
QY 1981 cattggaagcaagactgacactcctctctctctctctctctctctctctctctctctctct 2040  
Db 2014 CATTTGGACAGAGACTGACTCTCTTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2073  
QY 2041 aatgctctatgaaagacacactcaacccatctccatctcagaggaagaaactgtctcatgct 2100  
Db 2074 AATGCTTATGAAGACACACTCAACCTATTCATTCATTCATTCATTCATTCATTCATTCATTC 2133  
QY 2101 gatggaacccaggtctatgactctgtgggtgtccacaactcgaactcttggaacagagg 2160  
Db 2134 GATGGAACCAACCAAGCTTATGATGATTTCTGGGGTGCACACACTGACACTTTCGAAACAGAG 2193  
QY 2161 catgacgcctctgaaggttctctgattgtgacaagaacacatgagatattatagagga 2220  
Db 2194 CATGACCGCTTACTGAAAGTTTCTAGTTTATTCAGAGGGGAGAGAGACGACGACTA 2253  
QY 2221 caattatgaatatatca-----gcalactigtctgtaaaaaacaatgc 2265  
Db 2254 TCTGGACCTTGGAGAAATATTCAGTGAAGACGAGACTATATGACATCTGTCAGACTCT 2313  
QY 2266 catlgaacaaagaagctctcccaagaatlcagaacacccctagacactaggaagaagaatc 2325  
Db 2314 GATTGAACCAAGAAAGTCTTCCAGAAATTCAGAACACCCCTAGACATAGGCAAAAAGCAATT 2373  
QY 2326 taatgcaac----- 2334  
Db 2374 TAAATGCCACCAATTTCCAGAAATGACATAGAGAAAGACTGACCCCTTTGGTTGCACACAG 2433  
QY 2335 ----- 2334  
Db 2434 AACACCTATGCTTAAATAACAAATGTCCTCTAGTATTTGATTTGCTTTGGGACA 2493  
QY 2335 ----- 2334  
Db 2494 GAGTCTACTCCATGAGGCTATCTTATCTGATCTTCCAGAAAGCCAAATATGAGACTTT 2553

QY 2335 ----- 2334  
Db 2554 TTCTGATGATCATCACCCTGGAGCAATAGACAGTAATAACAGCCTGTCTGAATGACACA 2613  
QY 2335 ----- 2334  
Db 2614 CTTTCAGGCCACAGCTCCATCAGCTGGGACATGGTATTTAACCCCTGAGTCAGCCCTCCA 2673  
QY 2335 ----- 2334  
Db 2674 ATTAAGATTAAATGAAAGAACTGGGGACAACTGCAGATCCTCTTGCTGGGATTAACACATA 2733  
QY 2335 ----- 2334  
Db 2734 TGGTACTCAGATACCAAAAGAGAGTGGAAATCCCAAGAGAATCCACGAAAAAACAGC 2793  
QY 2335 ----- 2334  
Db 2794 TTTTAAAGAAAAAGATACCATTTTGTCCCTGAAACGCTTGTGAAGCAATGCATGCAATAGC 2853  
QY 2335 ----- 2334  
Db 2854 AGCAATTAATGAGGGAGACAAATTAAGCCCGAATAGAACTACCTGGGCAAGCAAGTAG 2913  
QY 2335 ----- 2334  
Db 2914 GACTGAAGGCTGTGCTCTCAAAAACCCACAGCTCTTGAACGCGCATCAACGGGAATTAAC 2973  
QY 2371 tctgtacctcttcagtcagatcaagaaggaaattgactatgatatacattatcagttga 2430  
Db 2974 TCGTACTACTCTTCAGTCAGATCAAGAGAAATTGACTATGATGATACCATATCATCTGTTGA 3033  
QY 2431 aatcgaagaaggaaatttgcacattatgatgatgaatgaataacacagagcccccagctt 2490  
Db 3034 AATGAAGAAGAGAGATTTTTCATTTATGATGAGAGATGAATAATCAGACCCCGCAGCTT 3093  
QY 2491 tcaagaagaacaacagacataattatctgtcagctgagagagctctgagatlatggat 2550  
Db 3094 TCAAAAAGAAAAACAGACACTATTTTATGCTGCAGTGCAGAGCGCTCTGGATTTATGGGAT 3153  
QY 2551 gattagctcccccacatgtctcaagaacaaggctcagagtgagtgccctccatgaataa 2610  
Db 3154 GAGTAGTCCCCACATGTCTTAAGAAACAGGCTCAGAGTGCGAGTGCCCTCACATTAA 3213  
QY 2611 gaaagtcttccagaagaatttactgatctgctcttactcaagccctataacgttgaaga 2670  
Db 3214 GAAAGTGTTCCTCAGGAATTTACTGATGGCTCCTTTATCAGCCCTTAATACCGTGAGA 3273  
QY 2671 actcaaatgaacaatttggagctcctggggccaataaagaagcagaagttgaagataat 2730  
Db 3274 ACTAAATGAACATTTGGGACTCCTGGGGCCATTAATTAAGACAGAAAGTTGAAGATTAATAT 3333  
QY 2731 catgtgaacttcaagaatacagagcctgtgcctatccctctatctatctaatctatttc 2790  
Db 3334 CATGGTAACCTTTCAGAAATGAGGCTCTCGTCCCTATTCTCTTAATCTTAACTTATTC 3393  
QY 2791 tctatggagaagatcaagagcaagagcagaacactagaaaaaacttgcgaagcctaatga 2850  
Db 3394 TTTATGAGGAGATCAGAGGACAGAGCAGAACCTTAAGAAAAAATTTGTCAAGCCCTAATGA 3453  
QY 2851 aaccacaacttacttcttggagaagtgcacaatcataatggcaccactaaagaatggattga 2910  
Db 3454 AACCAAAACTTACTTTTGGAAAGTGCAAACATCATATGCGACCCACTAAAGATGAGTTGA 3513  
QY 2911 ctgcacaagccttggtctattctctgtatgttgacctggaagaagatgtgcactaagcct 2970  
Db 3514 CTGCAAAAGCCTGGGCTTATTTCTGTGATGTGACCTGAAAAAAGATGTGCACCTAGGCT 3573  
QY 2971 gatttgagccctctcgtctgcacacataacacactgaaccccttcacttggaagaagtt 3030  
Db 3574 GATTTGAGCCCTTCTGTGTCGCCACACTAACACACTGAAACCTGCTCATTTGGAGACAAAGT 3633  
QY 3031 gacagtacaggaatttgcctgttcttcaaccacttctgatagaagcaaaaaagctgctact 3090

Db 3634 GACACTACAGGAATTTGCTGTGTTTTTACCATCTTTGATGAGACCAAAAGCTGTACTT 3693  
QY 3091 cactgaaaaatctggaagaagaactcagaggtcccttgcaatataccaatggaaatccac 3150  
Db 3694 CACTGAAATATGGAAGAAACTGAGGCTCCCTGCAATATCCAGATGGAGATCCAC 3753  
QY 3151 tttaaagaagatlatcgtctccatctgaatcaatcagtcataatgatacactctg 3210  
Db 3754 TTTTAAAGAGATATTCGCTTCATGCAATCAATGGCTACATATATGATACACTCTGG 3813  
QY 3211 cttagtaatgtcgaagatcaagaagatcgatgtatctctcagatcagctggcagcaatga 3270  
Db 3814 CTATGATGCTCAGATCAAGATGCGATGGTATCTGCTCAGCATGGCGCAATGA 3873  
QY 3271 aacatccatctatctatctcagtgagacatgttcaacttgtaagaaaaaagagagta 3330  
Db 3874 AAACATCCATTCTATTCATTTTCAGTGGACATGTCTTCACTGTAGGAAAAAGAGAGTA 3393  
QY 3331 laaatgacactgtacaaatcctcatalccaggtgttcttgagaagtggaatgttaccatc 3390  
Db 3934 TAAATGSCACTGTACATCTCTATCCAGGCTTTTGAGACAGTGAATGTTTACATC 3993  
QY 3391 caaagctggaatttggcgggtggaatgccttatctggcgaacatcatalcgtcgtgatatg 3450  
Db 3994 CAAGCTGGAATTTGGCGGTGGAATGCTTATTTGGGAGATCTTACATGCTGGATGAG 4053  
QY 3451 cagacttctcgtgtgtaagaacaataagtgtaagacccctctggaaatgtcctctgaca 3510  
Db 4054 CACACTTTTGTGTGTACACAAATAGTGTACACTCCCTGGGAATGCTTGTGACA 4113  
QY 3511 catlagagatttcaagatlaacagctcagagaacaatalggaacagtgggcccaagctgac 3570  
Db 4114 CATTTAGAGATTTTTCAGATTAACAGTTCAGACATATGAGACAGTGGGCCCAAGCTGGC 4173  
QY 3571 cagacttcaatctccggatcaatcaatgccttggaagcaaccaagagcccttcttgat 3630  
Db 4174 CAGACTTCATTAATTCGGGATCAATCAATGCTTGAGAGCAACAGGAGCCCTTTCTTGAT 4233  
QY 3631 caagtggaatctgttgacccaatgatatctcagggacaaagccaggtggccctga 3690  
Db 4234 CAAGTGTATCTGTTGGACCAATGATTTATTCAGGCAATCAAGGAGTCCCTGCA 4293  
QY 3691 gaagtlctcagccctcatalctcagtttatalcatatglatatgltctgataggaaaga 3750  
Db 4294 GAAGTTCACAGCCTCTCATCTCTCAGTTATCATGATGATAGTCTGTATGGAGAA 4353  
QY 3751 gtggcagactatcggaggaattccactggaaacttaaggtcttcttggcaatgga 3810  
Db 4354 GTGGCAGACTTATCGAGGAATTCACCTGGAACCTTAATGCTCTTTGGCAATGTGA 4413  
QY 3811 ttcatctgggataaacaacaatatctttaaaccctcaatatctgcgtacatccgtt 3870  
Db 4414 TTTCATCTGGGATTAACCAATATTTTAACTTCACCAATTAATGCTGTGATACATCCGTTT 4473  
QY 3871 gcaccaactcatalatagcatctgcagaactcttcgcatggagtgaatggcctgtgatt 3930  
Db 4474 GCACCAACCTCATATTATGACATTCGACAGACTCTTCGATGAGTGAATGGGCTGTGATTT 4533  
QY 3931 aaatagtltgacgaatgcgaattgggaatggagagtaaaagcaatatcagatgcacagattac 3990  
Db 4534 AAATAGTTGCGAGCATGCGCATTTGGAAATGAGATGAAGCAATATCAATCAATCAATCAAT 4593  
QY 3991 tgccttaacacttacttaacaaatagtttggcgaactgttcccttcaaaaagctgaactca 4050  
Db 4594 TGCTTCACTCTTATTTACCAATATATGTTTCCACCTGGTGTCTCTTCAAAAGCTGCATTTCA 4653  
QY 4051 cctccaaggagagtaatgcttggaagcctcagtgtaataatccaaagaagtgtgca 4110  
Db 4654 CCTCAAGGAGGAGTATGCTGTGAGACCTCAGAGTGAATTAATCCAAAGAGTGGCTGCA 4713  
QY 4111 agtggactccagaagaacaatgaaagtcagaagtaactactcaagggagtaaatcctc 4170

Db 4714 AGTGAGCTCCAGAGACAAATGAAGTCAACAGAGTAACACTACAGGAGTAATCTCT 4773  
QY 4171 gcttacagcatgatatgaaggagttccatctccatccagcaagtaagtgcacatgtg 4230  
Db 4774 GCTTACACGACATGATGTGAAGGAGTTCATCTCCAGCAAGTCAATGATGACATGAGT 4833  
QY 4231 gactctcttttcaagaatgcaagtaagtttctcaaggaaatcaagctccac 4290  
Db 4834 GACTCTCTTTTTCAGATGAGCAAGTAAGGTTTTCAGGGAATCAAGACTCTTTCAC 4893  
QY 4291 acctgtgtgaactcttagaccaccgttactgaactgactaccttgaatcaaccaca 4350  
Db 4894 ACCTGTGTGTAACCTCTTACAGCCACCGTTACTGACTGCTTACCTTGAAATCCACCCCA 4953  
QY 4351 gaattgtgtgacacagttgcccctgaagatgtaggttctgtgctgaggaagcaagacct 4410  
Db 4954 GAGTTGGGTGCACCAAGATTCCTCCAGAGATGAGAGTTCGGGTGCGAGGACAGACCT 5013  
QY 4411 ctactga 4417  
Db 5014 CTACTGA 5020

RESULT 10  
US-09-243-539-1  
; Sequence 1, Application US/09243539  
; Patent No. 6130203  
; GENERAL INFORMATION:  
; APPLICANT: VOORBERG, Johannes J.  
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/243,539  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/558,107  
; FILING DATE: 13-NOV-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: ISACSON, John P.  
; REGISTRATION NUMBER: 33,715  
; REFERENCE/DOCKET NUMBER: 30472/212  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5035 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 35..5017  
; US-09-243-539-1

Query Match 84.8%; Score 3743.8; DB 3; Length 5035;  
Best Local Similarity 87.5%; Pred. No. 0;  
Matches 4365; Conservative 0; Mismatches 52; Indels 570; Gaps 2;

QY 1 catgcaaatagagctctcaacgtctctcttctgtgaccttlttgagatctgcttttagtc 60  
Db 34 CATGGAATATAGAGCTCTCCACCTGCTTCTTGTGTGCTTTTGAGATTGCTTTAGTGC 93  
QY 61 caccagaagaatacctctggtgtgagtggaactgtcaatgggaactatataatgaatgta 120  
Db 94 CACGAGAAGATACACTCTGGGTGAGTGAAGTTCATATGGGACTATATGCAAGTGTCT 153  
QY 121 cggtagagctgtcgtgtgagcgaagatltcctcctagagtgccaaatccttccatcaa 180  
Db 154 CGGTGACCTGCTGTGAGCCCAAGATTTCCTCTAGAGTGCACAAATCTTTCCATTCAA 213  
QY 181 caactagctgtgtacaaaagaactcgtttgtagaattcaaggttcaacttltcaaat 240  
Db 214 CACCTCAGTGTGTACAAAAGACTGTTGTATGAATTCAGGATACCTTTTCAACAT 273  
QY 241 cgttaagcacaagcacaaccctgagatggtctgtagagctccaccatccagctgagttta 300  
Db 274 CGCTAACCCAAAGGCCACCCCTGGATGGGTCTGATAGTCTTACCATCCAGGCTGAGTTTA 333  
QY 301 tgatcacgtgtcaatlaacttaagaaacatggtcttccatcctgtcagttctatgtct 360  
Db 334 TGATACAGTGTGATATACCTTAAGAACATGGCTTCCATCTGTCTAGTCTCATGCTGT 393  
QY 361 tgggtgtatcctctctggaagaactctctgagggagctggaatagatagatcagaccag 420  
Db 394 TGGTGTATCCTACTGGAACCTTCTGAGGAGCTGGAATATGATATGACACCACTCAAG 453  
QY 421 gggagaagaagatgataaagcttccctgtgtgaagcatataltctcggcaggtcct 480  
Db 454 GGAAGAAAGATGATATAAGTCTTCCCTGTTGGAAGCCATACATATCTTGGCAGGTCT 513  
QY 481 gaaagaatgtgtccaaatgagcctctgacccactgtgcttactactacatacttctca 540  
Db 514 GAAAGAAATGTGTCCATATGAGCTCTGACCCAGTGTGCTTACCTACATATCTTCTCA 573  
QY 541 tgtgagactgtgtaaaagactgaattcagagcctcatatgagggccctactaataatgtaga 600  
Db 574 TGTGAGACTGTGTAATAAGCTTGAATTCAGGCTCATATGAGCCCTTACTATATGTAGA 633  
QY 601 aggaagctgtgccaagaaagacacagacctgtgacaaatataactacttltgtct 660  
Db 634 AGGAGCTGTGGCCAGSAAAGAACACAGACTTGCACAAATTTATACTTTTGTCTGT 693  
QY 661 atttgatgaaggaaagtgtgcaactcagaacaaagaactccttgaatgagatagga 720  
Db 694 ATTTGATGAAGGAAAGTGTGCACCTCAGAAACAAAGACTCTTGTATGAGATAGGA 753  
QY 721 tgcctgcatcgtcctcggcctggtcctaaatgacacagtaagttaataacaggtc 780  
Db 754 TGCCTGATCTGTCTGGGCTGTGAAATGACACAGTCAATGTATGTAAACAGTCT 813  
QY 781 tctgccaagctgcatgtagtgaaccaagaaatcagttatgtgcatgtgaatggg 840  
Db 814 TCTGCCAGGTCTGATTTGATGACCAAGAAATCAGTTATTTGGCATGTGATTTGAAATGG 873  
QY 841 caccactccttgaagtgtcactcaatlatcctcgaaggtcacacatltctgttgaagacca 900  
Db 874 CACCACTCCTGAAGTGCATCAATATTCCTGAAAGTGCACAAATTTCTGTGAGAACCA 933  
QY 901 tgcgcaaggtcctctgtgaatctcgccaataacttctcctactgctcacaactctgtat 960  
Db 934 TCGCCAGGCTCCTTGTGAAATCTCGCCAAATTAATCTTCTTAACTGCTCAACACTCTGTAT 993  
QY 961 ggaaccttgacagtttctactgttttgaatcactcttccaccacaatatgtgcatgga 1020  
Db 994 GGAACCTTGGACAGTTTCTACTGTTTGTCTATATCTCTTCCACCAACATATGAGCATGA 1053  
QY 1021 agcttatgtcaaatgtagacagctgtccagaggaaccaccaactagaatgaataatga 1080  
Db 1054 AGCTTATGTCAAAAGTGAAGAGCTGTGCAGAGAACCCCACTGAGATGAAAATATATGA 1113

|    |      |   |      |
|----|------|---|------|
| OY | 1081 | agaagcggaagactatgatatgatacttactgattcttccaatggaatggaatggttcgaagtttga | 1140 |
| Db | 1114 | AGAAGGGGAAGACTATGATGATGATGATCTTCACTGATTCTGGAATGGATGGTGGCAGGTTTGA    | 1173 |
| OY | 1141 | tga tga caacttccttccttccttccaaattgcgtcagtttgccagaagca tcccta aac    | 1200 |
| Db | 1174 | TGATGACMACTCTTCCTTCTTTATCCAAATTCCGTCAGTTGGCCAAGAACATCTTAAAC         | 1233 |
| OY | 1201 | tttggtfacattacatctgctgctgaagaggaagcttggagctatgctccctagtccctgc       | 1260 |
| Db | 1234 | TTGGGTACATTACATTGCTCTGCTGAAGAGAGACTGGGACTATGCTCCCTTAGCTCTGCG        | 1293 |
| OY | 1261 | ccccgctgacagaagatataaaagtcataatttgaacaatggcccccaagagatctgtag        | 1320 |
| Db | 1294 | CCCCGATGACAGAAGTATTAAGTCAATATTGTAACAAATGGCCCTCAGCGGATGGGTAG         | 1353 |
| OY | 1321 | gaagta caaaaagctccgattcttgaatacacagatcgaaccccttaagctctgtagc         | 1380 |
| Db | 1354 | GAAGTACAAAAAGTCCGATTTTATGGCATACAGATGAACCTTTAAGACTCGTGAAC            | 1413 |
| OY | 1381 | tattcagcatgaa tca gnaa lcttgygaccttacttlatygggaag tlgaga cacact     | 1440 |
| Db | 1414 | TATTCAGCATGATCAGGAATCTTGGACCTTTACTTTATGGGGAAGTTGGAGACACACT          | 1473 |
| OY | 1441 | gtttgatataatlaagaatcaagaagaacatataaactctaacctcaagaa tccac           | 1500 |
| Db | 1474 | GTTTGATTATATTTAAGATACCAACCAACCATATTAACATCTACCCCTCAGGAAATCAC         | 1533 |
| OY | 1501 | tga tgccttccttcttlatccaaggagatataccaaaagctglaaaacatttgaagattt       | 1560 |
| Db | 1534 | TGATGTCCTGCTTTTGATTCAAGGAAATTTACCAAAAGTGTAACATTTGAAGGATTT           | 1593 |
| OY | 1561 | lccaa tctgcagaga gnaa taltccaata laa tlgacag tga ctg laga aatlgggcc | 1620 |
| Db | 1594 | TCCAAATTCTGCAGGAGAAATATTCAATATTAATGACAGTGCATGACAGATGGGCC            | 1653 |
| OY | 1621 | aactaa tca gaatccctgcgtgctgacccgctatctactagtcttgtaata tggagag       | 1680 |
| Db | 1654 | AACATAATCAGATCTTCGGTCCGTGACCCTCTTACTTCAAGTTCCGTATATGAGAG            | 1713 |
| OY | 1681 | agatcagcttcaagagctaatlgggccctccccaatctgtacaaagaatctgtagata          | 1740 |
| Db | 1714 | AGATCTAGCTTCAGGACTCACTTGGCCCTCTCTCATCTGCTACAAANATCTGTAGATCA         | 1773 |
| OY | 1741 | aagaggaaacccagataatgtcagaacaagaggaa tgcatacctgtcttctglatlttga       | 1800 |
| Db | 1774 | AAGAGAAACCAAGATATATGTCAGACACAAGAGAAATGTCATCTCTGTATTTGATGA           | 1833 |
| OY | 1801 | gaacccaagctgtgacctccaagaagaatatataacgcttcttcccaca tccaagctgaat      | 1860 |
| Db | 1834 | GAACCCAACCTGGTACTCTCAAGAGATATTAACAAGCTTCTCCCAATCCAGCTGAGAG          | 1893 |
| OY | 1861 | gaagctctggaatccagagttccaaagctccaacatctgacagaa taaatgtcattc          | 1920 |
| Db | 1894 | GCAGCTTGAGAGATCCAGAGTTCCAAGCCTCCAACATATGCACAGCATCAATGGCTATGT        | 1953 |
| OY | 1921 | tttctgaatcttgcaatctgtcagtttggatctgcatlgagtgagca baactgtaacattcaag   | 1980 |
| Db | 1954 | TTTTGTATTTGCAATTGTGCATTTGTTGTCATGAGAGTGGCACTAGTATCAATTTAAG          | 2013 |
| OY | 1981 | catitgagacacagactgacttcccttctgtccttctctcgtgataatactccaacacaa        | 2040 |
| Db | 2014 | CATTGGAGACACACATGACTTCTTTCTGTCTTCTCTTCGTGGATATCTTTCAAAACAA          | 2073 |
| OY | 2041 | aatgtccttgaagaacaactcaacccattcccaattctgaagaaactgtcttcatgctc         | 2100 |
| Db | 2074 | AATGGCTTATGAAAGACACACTCAACCTATTTCCATTTCTCAGGAGAAACTGTCTTCAATGC      | 2133 |
| OY | 2101 | gactgaaaacccagctcctatggaattctgggttgccaacaactcagacttctggaacagag      | 2160 |
| Db | 2134 | GATGCAAAACCCAGGCTATGAGATTTCTGGGGTGGCAACATCAGACTTTGGGAACAGAG         | 2193 |
| OY | 2161 | catgacgccttactgaaggttctctagtttggacaagaacactgtgatatattacagaga        | 2220 |

|    |      |  |      |
|----|------|--|------|
| Db | 2194 | CATGACGGCCTTACTGAGAGGTTTCTAGTTGATTCCAGAGGGGGAGGAGACGACACTA       | 2255 |
| Oy | 2221 | cagttaagaagatalttca-----gcacttctgagtaaaacatgac                   | 2265 |
| Db | 2254 | TCTGGACCTGGAGAGAAATTCATCATGAAAGAGAGACATACATGACATGTCGACAGTC       | 2313 |
| Oy | 2266 | catgaaccaaagaagctctccagaaatcaagaacaccctagcataggcaaaagaat         | 2325 |
| Db | 2314 | GATTGAAACCAAGAGGCTTCTCCGAAATTCAGAGACCCCTAGCAGTACGCAAAACCAATT     | 2372 |
| Oy | 2326 | taatgcacac-----  | 2334 |
| Db | 2374 | TAAATCCACCAACAAATTCAGAAAAATGACATAGAGAGACTGACCTTGCTTTCACACAG      | 2433 |
| Oy | 2335 | -----  | 2334 |
| Db | 2434 | AACACCTATGGCTAAATACAAAATGTCCTCTAGTAGATTGTGTATGCTCTGTGGACA        | 2499 |
| Oy | 2335 | -----  | 2334 |
| Db | 2494 | GAGTCCTACTCCACATGGCGCTATCCTTATCTGATCTCCAGAAAGCCAAATATGACACTT     | 2553 |
| Oy | 2335 | -----  | 2334 |
| Db | 2554 | TTCTGATCATCTCCATCCACTCGAGACAAATAGACAGTAATACAGCCTTCTGAAATACACA    | 2613 |
| Oy | 2335 | -----  | 2334 |
| Db | 2614 | CTTCAGGCGCACAGCTCCATCCATCAGATGGGGACATGTGATTATCCCTTAGCTACAGGCTCCA | 2673 |
| Oy | 2335 | -----  | 2334 |
| Db | 2674 | ATTAGATTAAATGAGAAACTGGGGACACTGCAGATCCTCTTGGATTGGAATACACTA        | 2733 |
| Oy | 2335 | -----  | 2334 |
| Db | 2734 | TGCTACTCAGATACCAAAAAGAGAGTGGAATCCCAAGAGAGCTCACGAAAAAACAGC        | 2793 |
| Oy | 2335 | -----  | 2334 |
| Db | 2794 | TTTTAGAAAAAGCATACATTTTGTCCCTGAAGCCTGTGGAAGACATATGCATATGC         | 2855 |
| Oy | 2335 | -----  | 2334 |
| Db | 2854 | AGCAATAATGAGGACAAAAATAGCCCGAAATAGAAATCACCCTGGCAAAACAAGTAG        | 2913 |
| Oy | 2335 | -----  | 2370 |
| Db | 2914 | GACTGAAAGGCTGTGCTCTCAAAACCCACAGCTTTGAAAACGCAATCAACGGGAAATAAC     | 2973 |
| Oy | 2371 | tcgtactactcttcagtcagaatcaagaagaaatctgactatgatatcatatcagttga      | 2430 |
| Db | 2974 | TCTGACTACTCTTCTCAGTCAGATCAAGAGAAATTGACATGATGATACCATATTCAGTTGA    | 3033 |
| Oy | 2431 | aatgaagaagaagatttgacattatgatgatgaatgaatcaagaaccccgacgtc          | 2490 |
| Db | 3034 | AATGAAAGAGGAATATTTTGACATTTATGATGAGGATGAAAAATCAAGAGCCCGCAGCTT     | 3093 |
| Oy | 2491 | tcaaaagaagaacagaacataatttatctgtagtggagagagctctgtagtlatagagat     | 2550 |
| Db | 3094 | TCAAAAGAAAACGACACTATTTTATTTATGCTGAGTGGAGAGGCTCTTGGGATTAATGGAT    | 3153 |
| Oy | 2551 | gagtagcccccacatgcttctlaagaaacagagctcagagtygagagtgctccctcagttcaa  | 2610 |
| Db | 3154 | GAGTAGCTCCCCACATGCTTCTAAGAAACAGAGGCTCAGAGTGGGAGTGTCCCTCACTTCAA   | 3213 |
| Oy | 2611 | gaaagtgttttccagaatattactgaatgctccttactcagcccttataccgtlgsaga      | 2670 |
| Db | 3214 | GAAAGTTGTTTTCAGGAAATTTACTGATGCTCTTACTACAGCCCTTATACCGTGAGAG       | 3273 |
| Oy | 2671 | actaaatgaacatttggactcctctggggccatatataagagcagaagtgtgaagataatat   | 2730 |

Db 3274 ACTAATGACATTTGGGACTCTGGGGCCATATATAGACAGAGTGAAGATAATAT 3333  
OY 2721 catgtaacttcaagaatacagcctctcctcattccttcaatccttattc 2790  
Db 3334 CATGTAATCTTCAAGAAATCAGGCTCTCGTCCATTCTCTATTTCTACCTTATTTTC 3393  
OY 2791 tcatggaagaatcagaagcagaagcagaacactagaanaaacttctgaagcctaa 2850  
Db 3394 TTATAGAGAAATCAGAGGAGAGAGAGAAACCTAGAAAACCTTTGCAAGCCTATATGA 3453  
OY 2851 aaccaaacttacttcttggaagtgcaacatcatactgaccctaaagaatga 2910  
Db 3454 AACCAAAACTACTTCTTGGAAAGTGCAACATCATATGGACCCACTAAAGATGATTGGA 3513  
OY 2911 ctgaaagccggcttattctcttgatggtgaccttggaagaatgctgacctgaagcct 2970  
Db 3514 CTGCAAGGCTGGGCTTATTTCTGTGATGTGACCTGAAAAAATGTGCACTGAGCCT 3573  
OY 2971 gattggaacccctctgctcgcacactaacacatgaacctgctcatggaagcaagt 3030  
Db 3574 GATTGACCCCTTCTGCTTCGCCACTAACACACTGACCTCTCATGGAGACAAAGT 3633  
OY 3031 gacagtaacgaagatctgctcgttcttcaacatctctgtagagaccaaagctgtaact 3090  
Db 3634 GACAGTACAGGAATTTGCTGTTTTCACCATCTTGATGAGACCAAAAGCTGTAATT 3693  
OY 3091 cactgaaataatggaagaacatgcaagcctcctgcaataatcagaatggaagctccac 3150  
Db 3694 CACTGAAATATGGAAGAACTGCAAGGCTCCTGCAATATCCAGATGGAAGTCCAC 3753  
OY 3151 tttaaaagagaatatacgtctcctcaatgaatgctacataatgatacaactctg 3210  
Db 3754 TTTTAAAGAGATATATGCTTCATGCAATCAATGCGTACATATGATACACTGCG 3813  
OY 3211 cctagtaatggtcagaatcaaaagatctgctgctcctgcaagctggaagcaatga 3270  
Db 3814 CTGTGATATGCTCAGATCAAAAGATTCGATGCTATCTGTCAGCTGGGCAACAATGA 3873  
OY 3271 aaacatcatcttaattcaatcagtgagatggttcaactgtaagaanaaagggagta 3330  
Db 3874 AAACATCATTTCTATTTCAATGAGACATGCTGTCAGTACGTAAGAAAAGAGAGTA 3933  
OY 3331 taaaatgacactgacatcctcctcaatcaggtcttcttgagaagctggaatctaccat 3390  
Db 3934 TAAATGGAAGTACATCTCTATCCAGGCTTTTGAAGACATGGAATGTTACCATC 3993  
OY 3391 caaagctggaatttggcgggttggaatgcttaattgagcagcatctacatgctggatga 3450  
Db 3994 CAAAGCTGGAATTTGGCGGGTGGAATGCTTATGCGAGCATCTACATGCTGGATGAG 4053  
OY 3451 cagactttctctgtgtaacagcaataagtgctcagactcccttgggaatgctcttgaca 3510  
Db 4054 CACACTTTTCTGCTGTAAGCAATTAAGTGTGACAGTCCCTGGGAATGGCTTTTGACA 4113  
OY 3511 catagaagatttcaagattcagctcagaacaaatcagcaagtgagggcccaaaagctggc 3570  
Db 4114 CATTAGAGATTTTCAGATTCAGCTTACAGACATATGAGACAGTGGGCCCAAAAGCTGGC 4173  
OY 3571 cagactcatattcggatcaatcaatgctcctgagaccaaagagcccttcttgat 3630  
Db 4174 CAGACTTCATTAATCCGATCAATCAATGCTGAGCACCAAGAGGCCCTTTCTTGAT 4233  
OY 3631 caaagtgaatctcttgacacaaatgattatcagcgaataagaaccaaaggtgcccgtca 3690  
Db 4234 CAAAGTGAATCTGTTGGACCAAGATTAATTCAGGCATCAAGGCCAGGGTGCCTGCA 4293  
OY 3691 gaagctcagcctcactacatctcagttatcatcatglatatgcttctgtagaagaa 3750  
Db 4294 GAAATTTCCAGCCTCTACATCTCTCAGTTATATCATGTATGTTGATGGAGAA 4353  
OY 3751 gtggcagacttatcagagaaatccactggaaccttaatggcttcttctggcaatgltga 3810  
Db 4354 GTGGCAGACTTATCGAGAAATTCACCTGGAACCTTAATGCTCTTCTTGGCAATGAGA 4413

OY 3811 ttcatctgggataaaacaaatattttaaaccctccaattatgctcgatatacgcgttt 3870  
Db 4414 TTCTATCTGGGATTAATAACATATATTTTAACCTTCATATTTGCTGATACATCCGTTT 4473  
OY 3871 gcaccacactcatatagatctcgcaagcactcttcgcatgagtgatgagtggtctgatt 3930  
Db 4474 GCACCAAGTCATATATGATGATTCGACACACTCTCGCATGAGTGAATGGGCTGTGATTTT 4533  
OY 3931 aaatagttcagcatgcatggaatggaatggaatgaagaatcagatgcaagatctac 3990  
Db 4534 AAATAGTTGCAAGCATGTCATTTGGAAATGAGAGATGAACATATCAATGCTCACGATTAC 4593  
OY 3991 tgcctcaactcttcaacaataatggttgcaacctggtctccctcaaaagctgactca 4050  
Db 4594 TGCTTCATCTTCTTACATATATGTTTGGCCACTGGTCTCTCAAAAGCTCGACTTCA 4653  
OY 4051 cctccaagagagagatgactcctgagacactcaagtgataatccaagaagtgcgtga 4110  
Db 4654 CTTCAAGGAGAGATATGCTGCTGACACCTGACAGTGAATATCCAAAGATGCTGCA 4713  
OY 4111 agtggacttccagaagaacaatgaagtcacagagtaactactaaggagtaaaatctct 4170  
Db 4714 AGTGACTTCCAGAAAGCAATGAAGTCAAGAGATTAATCACTGAGAGTAAATCTCT 4773  
OY 4171 gcttaccagatgatgtggaagaggttcctcatctcagaagatgaatgccaatgag 4230  
Db 4774 GCTTACCAAGCATGTATGTAAGAGGTTCTCATCTTCCAGCATGAATGAGTGCATGAG 4833  
OY 4231 gactccttcttcaagaatgcaagtaaggttctcagaagaaatcaagctcctcac 4290  
Db 4834 GACTCTCTTTTTCAGAAATGGAAGTAAAGGTTTTCAGGGAATCAAGCTCTTTCAC 4893  
OY 4291 aactgtgtgaactcctcagaacccaacggttaactgactccttgaatcaaccacca 4350  
Db 4894 ACCGTGTTGAACCTCTAGACCCACCGTACTGACTGCTTGAATTCACCCCA 4953  
OY 4351 gaattgggtgacacagattccctcagagatggaaggtcttgagctggcagagacagagact 4410  
Db 4954 GAGTGGGTGACACCAATGCTGCTGAGAGATGAGGTTCTGGGCTCGAGGACAGGACT 5013  
OY 4411 ctactga 4417  
Db 5014 CTAAGTA 5020

RESULT 11  
US-08-670-707A-38  
; Sequence 38, Application US/08670707A  
; Patent No. 5859204  
; GENERAL INFORMATION:  
; APPLICANT: Lollar, John S.  
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Circle Suite 201  
; CITY: Boulder  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/670,707A  
; FILING DATE: 26-JUN-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/US94/13200  
; FILING DATE: 15-NOV-1994







|    |      |  |       |
|----|------|--|-------|
| Db | 1680 | AAAGATCTGCTTGGGACATTTGGCCCTCTCTCATCTGCTACAAAGAATCTGTAAAC           | 17399 |
| OY | 1739 | caaaaggaaccagaaataatgtcacagaagaagtatcatctgttcttcttattat            | 17988 |
| Db | 1740 | CAAAAGAGAAACCGATGATGTACAGACAGAGAAACGTATCTCTTTTCTGTATTGCAT          | 17999 |
| OY | 1799 | gagaaacgaagctgtgtactctcacagagaataataacgcttctccccaatcagcttga        | 18588 |
| Db | 1800 | GAGATCAAAAGCTGTGTACCTCCGACGAGAAATATTACACGCTTCTCCCAATCCGATGGA       | 18599 |
| OY | 1859 | gtgagacttgaagatccaaggtttcaagctccaagctccaatcatatgcaacagatcaatgtgcat | 19188 |
| Db | 1860 | TTACAGCCCCAGGATCCAGAGTTCACAGCTTCTAACATATATCACAGATCAATGTGCTAT       | 19199 |
| OY | 1919 | gttttgaatgttgcagctgtgtcaagtttgttgcattgaagttgcaactgtgttaactta       | 19788 |
| Db | 1920 | GTTTTTGATACCTTTCAGCTGTGTGGTTTTGTGTTCACAGAGGTGTGCACTGTGATCTTCTA     | 19799 |
| OY | 1979 | agcatgtgagacacagacttgaacttcttctbcttcttcttcttctgtgatalaacttccaaac   | 20388 |
| Db | 1980 | AGTGTGGACAGACAGAGGAGCTTCTCTCCGTCGTCCTTCTGTGGCTAACACCTTCAACAC       | 20399 |
| OY | 2039 | aaaaatgtcatatgaagacaacaccacacctatcccatcttcagaggaactgttctcatg       | 20988 |
| Db | 2040 | AAAATGGCTATGTAAGACACACACACACCCGTTCCCTTCACGAGAAACCGCTTCTATG         | 20999 |
| OY | 2099 | tcgaatggaacccaagatctatgtacttctcgggggtgcacaactcagacttctcgaaaca      | 21588 |
| Db | 2100 | TCAATGAGAAACCCAGGTCTCTGTGGTCTTAAGGTGTGCACAACTCAGACTTGTGGAGACGA     | 21599 |
| OY | 2159 | ggacatgacgccttactbtaagagtttctagttgtgtgaacaagaacactgtgtatattacag    | 22188 |
| Db | 2160 | GGGATGACACACCTTACGTAAGGTGTATAGTGTGACAGGACGACATTTGTGATTAATTATGAC    | 22199 |
| OY | 2219 | gaacgtatataagatatlttcagcatactgtctgtgagttaaaaaatgtccatttgaaccaaga   | 22788 |
| Db | 2220 | AACACTTATGAAGAATATTTCCAGGCTTCTTGTGAGTGAAGAAAGTCAATTGAACCCAGA       | 22799 |
| OY | 2279 | agcttccccaagatttcaagaacacctagcactlagaacaaagcaatttaatgtcaacccca     | 23388 |
| Db | 2280 | -----  | 2279  |
| OY | 2339 | ccagltcttgaacgcatacaacgynaabaactcgttactactccttcagttcacatcaagag     | 23988 |
| Db | 2280 | -----GACATTAAGCGTTCTTACTTTTCACGCCGAGAGGAAGAC                       | 2315  |
| OY | 2339 | gaaatgtactatgatatatacatatcagtttgaatagaagaagaatttgaacattat          | 24588 |
| Db | 2316 | AAAATGGACATATGATATATCTTCTCACTCACTGAAGAAAGGAGAAATTTTGACATTTTAC      | 2375  |
| OY | 2459 | gataagaatgaabaalcaagagccccccgacgcttcaaaaagaacacgacatattttat        | 2518  |
| Db | 2376 | GGTGAAGATTAATAATACGAGACCTCCGACCTTTTCAGAAAGAACCCGACACTTAATTTCAT     | 2435  |
| OY | 2519 | gcttcagatvgagaagctctgtgatatatvggaatvgatlaagctcccaacagttcttaagaac   | 2578  |
| Db | 2436 | GCTCGGTGAGACAGGCTGTGGATATACGGGATGAGAGAAATCCCCCGGGCCCTTAAGAAC       | 2495  |
| OY | 2579 | agggatcaagatvggcaggtgtccctcaagtccaagaagaagtgttcttccaggaattactgat   | 2638  |
| Db | 2496 | AGGGCTTCAGAACGGAAGGTGCTCGTTCGTTCAAGAGGTGTGTTCCGGGAATTTGTCTGAC      | 2555  |
| OY | 2639 | ggaccccttaactaagccctataacacgttgaagaactaaatgaacatttbgagctctggg      | 2698  |
| Db | 2556 | GGCTCTTTCACGAGCCGCTGTACCCCGGGGAACCTAACAAACACTTGTGGGCTCTTGGA        | 2615  |
| OY | 2699 | ccatatataagagcagaagttctgaagataatcatcatgttaacttccagaatccaagccct     | 2758  |
| Db | 2616 | CCCTACATCAAGCGGAAGTGTGAAGACAACATCATGTGTACTTTCAAAAACAGCGCT          | 2675  |
| OY | 2759 | cgctccattctccttacttactagccttatttcttattgagaagatcagagcagaagca        | 2818  |

|    |      |  |       |
|----|------|--|-------|
| Db | 2676 | CGTCCCTATTTCCTTCTCTACTGAGCCCTTATTCTTATTCGGATGATCAGGACCAAGGCGCA     | 27353 |
| QY | 2819 | gaaccctagaanaaaactcttgclocaagcctaatagaaaccaaaacttactcttggaagtgcaaa | 28789 |
| Db | 2736 | GAACCTGCAGACACACTTGTGTGTACAGCCAAATGAAACCAAGAACTTACTTTTGGAAATGTGCAG | 27950 |
| QY | 2879 | calcatabvgcaccocctaagaabtgabtttgaactgcnaaagcctggagctatttcctgat     | 29388 |
| Db | 2796 | CATCATGAGCGCACCCACAGAAAGACGATTTTGATCTGCAGAAAGCCCTGGGCCCTACTTTCGTAT | 28555 |
| QY | 2939 | gttgaacgggaanaagttgagcactaagcgcgtgatgtggacccctctgctcgcaacact       | 29988 |
| Db | 2856 | GTTCAGCTGGAAAAAGATGTCTCATCGCTTATGTATGGCCCCCTCTGTATCTGCGCGCCGC      | 29155 |
| QY | 2999 | aacacactgnaacccctgtctcaltg99agaacaagtacagaacgaattgctcgtgtcttc      | 30588 |
| Db | 2916 | AACACCCCGAAGCGCTGTCTCAACGGTAGACAAAGTACCCGTGCAGAAATTTGCTGTGTTTTTC   | 29757 |
| QY | 3059 | aacaccltttgaatgagacccaagaagctgtgtacttcaactgnaaataatgynaagaactgcaag | 31188 |
| Db | 2976 | ACTATTTTTTGATGTAGACAAAGACGCTGTACTCTCATGAAAAATGTGAAAGAACTGCGCGG     | 30355 |
| QY | 3119 | gtccctctgcaataltccagatggaagaatccocactttaagaagaattatgccttcatabgca   | 31788 |
| Db | 3036 | GCCCCCTCCACCTGTGAGATGTGAGAGACCCCACTCTAAAGAAAACTATCGCTTCCATCA       | 30950 |
| QY | 3179 | atcaatgtgctacataatlygatacactaccctggttagtaatgctccagaatccaagaagat    | 32388 |
| Db | 3096 | ATCATGTGCTTATGTGATGAGTACACTCCCTGGCTTATGTATGGCTCAGAAATCMAAGGATC     | 31355 |
| QY | 3239 | cgaatgatactcgtcccaagcatlyggcgaagcaalyaanaacatccactctatcattcactgag  | 32988 |
| Db | 3156 | CGATGTGTATCTGTCTCACCATGTGGCGACAGAAATGAAATATCCATTGATTCATTTTAGCGGA   | 32155 |
| QY | 3259 | catgtytctcacttgcatacgaanaaagaagagtataaaaatgcaactgtaacatctctatcca   | 33588 |
| Db | 3216 | CACGTTTCACCTGTGACGGAAAAAGAGAGATTAATAAATGGCGGTACATATCTCATATCGG      | 32757 |
| QY | 3359 | ggtgtcttttggaagaaglygaatlygttaccatcccaagcctggaattctggcgggtggaatgct | 34188 |
| Db | 3276 | GGTGTCTTTGAGACAGTGGAAATGTCTAACCTGCCAAAGTTTGGAAATTTGGCGAATAGAAATGC  | 33355 |
| QY | 3419 | cttatlygcgaagcaltacalagcttggatlyagcacacttttctgtytgcacgaataag       | 34788 |
| Db | 3336 | CTGATTTGGCGGACACTGTCAAGCTGGGATGTAGACACACTTTTCTGTGTACAGCAAGGAG      | 33950 |
| QY | 3479 | tgtaagactccccctgggaatlyggtctctgcgaacaattagaagatltccaagatlaacagctca | 35388 |
| Db | 3396 | TGTAGGCTTCACAGTGGGAATGTGGCTTCTGTGACGGATTAAGATTTTTCAGATCTCACACTTCA  | 34555 |
| QY | 3539 | ggacataatgnaaagtggggccccaagaagctbvgccaagctcatlaltccagatccaatcaat   | 35988 |
| Db | 3456 | GGACAGTATGACAGTGTGGGCCCAAAAGCTGTGCGAGCTTCATTATTTCCGATCAATCATAT     | 35155 |
| QY | 3599 | gctctgagacacaagaagacccttcttcttggatcaaggtgagatctgltgcgaaccaatgata   | 36588 |
| Db | 3516 | GCTGTGAGACACAAGAGATCCCACTCCTGTGATCAAGTGTGATCTGTGTGGCACCAATGTATC    | 35757 |
| QY | 3659 | attcaagggatcaagaaccaaggggtgcgcgtcaagaagttctccaagctctacatctctcaag   | 37188 |
| Db | 3576 | ATTTCAGGCAATCAAGACCCAGGGGTGCCGTGACGAAGATTTTTCAGGCTCTACATCTCCAG     | 36355 |
| QY | 3719 | tttctacatcagtaatgltcttgaatgggaagaagtggccaactctacaggggaatctcaact    | 37788 |
| Db | 3656 | TTTATCTCATGTATCACTCTTTTGACGGGAGAGAACTGTGSCAAATTTACCGAAGGAAATTCACG  | 36950 |
| QY | 3779 | ggaaccttlaatgctctctcttctgcaaatgtgatactcaatcgggataaaacacataattctt   | 38388 |
| Db | 3696 | GGCACTTAATAGGTCTCTTTTGGCAATGTGTGAGACGATCTGTGGGATTTAAACACATATTTTT   | 37555 |
| QY | 3839 | aaacctccaattatgctctgatacatcogtlttgacaaccaactcatlatagatctcgaagc     | 38988 |
| Db | 3756 | AAACCTTCGATTTGTGGCTGTGGTACATTCCTTTTGACACCAACATTAATAGCATCGGACGC     | 38155 |

|    |      |             |              |         |           |        |          |        |                 |            |                  |        |        |        |      |      |      |      |     |     |      |
|----|------|-------------|--------------|---------|-----------|--------|----------|--------|-----------------|------------|------------------|--------|--------|--------|------|------|------|------|-----|-----|------|
| QY | 3899 | actcttcgcga | ttggaagttgat | tgggctg | tgatcttga | atlaa  | agtttcga | catgc  | atcttgggaatg    | 3958       |                  |        |        |        |      |      |      |      |     |     |      |
|    |      |             |              |         |           |        |          |        |                 |            |                  |        |        |        |      |      |      |      |     |     |      |
| Db | 3816 | ACTCTTCGCAT | GGAGTTGAT    | GGGCTG  | TGATTTAA  | ACAG   | TTTGAC   | ATG    | CCCTGGGAATG     | 3875       |                  |        |        |        |      |      |      |      |     |     |      |
| QY | 3959 | gagagtaaa   | gcaatc       | atcaga  | tgcag     | aatat  | ta       | cttgc  | tta             | ccaatg     | 4018             |        |        |        |      |      |      |      |     |     |      |
|    |      |             |              |         |           |        |          |        |                 |            |                  |        |        |        |      |      |      |      |     |     |      |
| Db | 3876 | CAGAAATAAG  | CGATAC       | AGACTCA | CACAT     | CATAC  | GGCCCTC  | CTCCAC | CTAAGCAATATATTT | 3935       |                  |        |        |        |      |      |      |      |     |     |      |
| QY | 4019 | ggccacttg   | cttcctt      | ccaaga  | agcttg    | atctt  | ca       | acttcc | caagggag        | gagla      | 4078             |        |        |        |      |      |      |      |     |     |      |
|    |      |             |              |         |           |        |          |        |                 |            |                  |        |        |        |      |      |      |      |     |     |      |
| Db | 3936 | GGCACC      | TGGTCTCT     | CTTCA   | CAGCCCG   | ACTTC  | ACTCC    | AGG66  | CGAGAA          | TGCTTGCGGA | 3995             |        |        |        |      |      |      |      |     |     |      |
| QY | 4079 | cctcagg     | tgaata       | atccaa  | aagag     | tggct  | tg       | caagt  | tg              | gacttcc    | agaagaca         | 4138   |        |        |      |      |      |      |     |     |      |
|    |      |             |              |         |           |        |          |        |                 |            |                  |        |        |        |      |      |      |      |     |     |      |
| Db | 3996 | CCCCGGT     | GAGCAG       | CCAG    | AGAG      | TGGCTC | GAGTG    | GA     | CC              | TCGAGA     | GAAGAGCGTGAAGTTC | 4055   |        |        |      |      |      |      |     |     |      |
| QY | 4139 | acaaga      | gtta         | ctac    | tca       | aggag  | ta       | aaatc  | tc              | tgctta     | ccaag            | catg   | 4198   |        |      |      |      |      |     |     |      |
|    |      |             |              |         |           |        |          |        |                 |            |                  |        |        |        |      |      |      |      |     |     |      |
| Db | 4056 | ACAGGCAT    | CACCA        | CCAGGCG | GTGAAG    | TCCCTG | CTCAG    | CA     | GTATG           | GTAAAG     | ATTC             |        | 4115   |        |      |      |      |      |     |     |      |
| QY | 4199 | cccatctc    | agcag        | ta      | aaga      | tggc   | catc     | ag     | tgg             | atc        | tctt             | ttt    | ggaat  | 4258   |      |      |      |      |     |     |      |
|    |      |             |              |         |           |        |          |        |                 |            |                  |        |        |        |      |      |      |      |     |     |      |
| Db | 4116 | CTCGTGT     | CCAGT        | GTAG    | TAGAG     | AGCGCC | CGCTG    | GA     | CCCTG           | TTTCTT     | TTCAG            | AGAGCG | CACAGC | 4175   |      |      |      |      |     |     |      |
| QY | 4255 | aaggtt      | ttt          | tcagg   | aaatc     | aaagc  | cttc     | ctt    | ca              | ac         | ctg              | tg     | glaa   | ctct   | ctag | acca | acog | 4318 |     |     |      |
|    |      |             |              |         |           |        |          |        |                 |            |                  |        |        |        |      |      |      |      |     |     |      |
| Db | 4176 | AAGGTTTT    | CAGCG        | CATAC   | GGACT     | CTCTCC | AC       | CC     | CC              | CGTGTGA    | ACCGCT           | CTTG   | AG     | CCCCCG |      |      |      | 4235 |     |     |      |
| QY | 4319 | ttac        | pac          | cgc     | gtac      | ctt    | tg       | aat    | ctt             | ca         | cccc             | agag   | tgg    | tg     | ca   | ccag | atg  | gc   | ctg | ag  | 4378 |
|    |      |             |              |         |           |        |          |        |                 |            |                  |        |        |        |      |      |      |      |     |     |      |
| Db | 4236 | CTGTTAC     | CGCGG        | CTAC    | CTTG      | AGAT   | TC       | CAC    | CC              | CA         | CG               | AG     | CTG    | GGCG   | CG   | CA   | GTG  | CC   | CTG | AGG | 4295 |
| QY | 4379 | atcga       | gg           | t       | ct        | tg     | gc       | ct     | gc              | ag         | gc               | ag     | ac     | gc     | at   | ct   | ta   | ct   | ta  | gc  | 4417 |
|    |      |             |              |         |           |        |          |        |                 |            |                  |        |        |        |      |      |      |      |     |     |      |
| Db | 4296 | CTCGAG      | TTCT         | AG      | ATG       | TG     | AG       | CG     | CA              | CG         | AG               | AT     | CT     | CT     | ACT  | GA   |      |      |     |     | 4334 |

```

?      REGISTRATION NUMBER: 33, 878
?      REFERENCE/DOCKET NUMBER: 75-95F
?      TELECOMMUNICATION INFORMATION:
?      TELEPHONE: 303/499-8080
?      TELEFAX: 303/499-8089
?      INFORMATION FOR SEQ ID NO: 38:
?      SEQUENCE CHARACTERISTICS:
?          LENGTH: 4334 base pairs
?          TYPE: nucleic acid
?          STRANDEDNESS: double
?          TOPOLOGY: not relevant
?      MOLECULE TYPE: cDNA TO mRNA
?      HYPOTHEetical: NO
?      ORIGINAL SOURCE:
?          INDIVIDUAL ISOLATE: Factor VIII lacking B domain
?      FEATURE:
?          NAME/KEY: CDS
?          LOCATION: 3..4334
?      US-09-037-601-38

```

|                       |              |               |                 |              |
|-----------------------|--------------|---------------|-----------------|--------------|
| Query Match           | 72.2%;       | Score 3189.8; | DB 4,           | Length 4334; |
| Best Local Similarity | 83.5%;       | Pred. No. 0;  |                 |              |
| Matches 3692;         | Conservative | 0;            | Mismatches 637; | Indels 90;   |
|                       |              |               |                 | Gaps 3;      |

Matches 3692; Conservative 0; Mismatches 637; Indels 90; Gaps 3;

2 atgcaaatagagctctccaccgctcttcttcgtgcctttgcgattcctcgtcttaagtc 61

Db 3 ATGCAGCTAGACCTCTCCACCTGTGTCTTTCTGTGTCTTGCACACTCGGCTTTAGTGC 62

D6  
07 acagaagatcctacccggtgcag tggaaactgcacagggaataatgcagaatga---t 118  
|||  
63 ATCAGGACTACTCCCTGGCGCAGTGTCTCTGGGACTACCGGCAAATGAATCTC 122

119 ctcggtagactgtcctgtgtgacgcaagattctctctctagagatgccaanaatctttccattc 178

|    |     |   |     |
|----|-----|---|-----|
| Db | 123 | CTCCGTGAGCTGCACGTTGACACCAAGATTCTCTGCTACAGCGCCAGGAGCTCTTCCGTTG | 182 |
|----|-----|---|-----|

**DY** 179 aacaccctcagtcgtgtacaanaagaactcgtttgttagaatccaogtltcacctttcaac 238  
-----  
**DH** 183 gccctccgcacgcccctcgatcaaaagacgctcccttgccggacacccctccaggacgac 242  
-----

239 atcgcctaagccaagccacccctgatactgctagtctcctacatccagctgagtc 298

Db 243 GTTGCCAGGCCAGGCCACCATGGATGGGTCTGCTGGGTCTTACCATTCACAGCTGAGTT 302

QY 299 tatgatacagtggtcatctaacctaagaacatggcttcccatcgtagtcattcatgct 358  
|| || || || || || || || || || || || || || || || || || || || || || || ||

359 attaatatattctactgaaaaactctgaaagaaactgaatataatgaacacacactca 418  
Db 303 TAGGACACAGGCTGGTCTGTACCCCTGAAGAACATGGCTTCTCATCCCGGTAGTCTTCACGCT 362

Db 363 GTGGCGTCTCTTGGAAATCTTCCGAAGCGCGTGAATATGAGATTCACACCGCAA 422

419 agggagaagaagatgatataagtcctccctggtggaagccatacatatgtctgcaagtc 478

Db 423 A G G G A A G A G A C G A T A A A G T C C T T C C C G G T A A A A G C C A A A C C T A C G T C T G G C A G G T C 482

|    |     |   |      |
|----|-----|---|------|
| Db | 483 | CGAAGAAATGGTCCACAGCCTCGAACCCACCATTGTCTCACTCATATCTGTC      | 5422 |
| QY | 4/9 | cgaagaagaaatggtccaatggcttcgaacccacttgccttaactcatcatattttc | 538  |

539 catgtgacctgttaaaagactgaattcagcgccatcttgagccctactagatgtaga 598

Db 543 CACGTGACCTGGTGAAGACCTGAATTCGGGGCCCTCATTTGGAGCCCTGCTGTTGTAGA 602

QY 599 gaaaggagatctgtgccaagaagaacacagacacctgtgcacaaatttactactcttltgtc 658

659 qatlttgatgaagqaaaaatlltgcactcagaanaacaaagaactcctllgatgcagatag 718  
 DD GAAAGGAGATCTGACCAAGAAAGAGACCCAGAACCTTCACCGAAATTTGTACTACTTTTGTCT 862  
 803

Db 663 GCTTTGATGAAGGAAGTTGGCAGTCAGCAAGAAATGATCTCTGGACACAGGGCCATG 722

|    |      |  |      |
|----|------|--|------|
| OY | 719  | gagctgcacatctgctctgggcctggcctaaatgacacacgtaacgtgtatgtaacagg      | 778  |
| Db | 723  | GATCCCCCACCCTCCACGGGGCCAGCCTGGCAATGCGACACAGTCATAAGGCTATGTCACAGG  | 782  |
| OY | 779  | tcctctgcagatctgattgattgacacacgaagaatcaacttatctgtgcatgtatggagatg  | 838  |
| Db | 783  | TCCTCGCAGGCTTGATCGGATGTCTATAAGAAATACAGTCATCTGACACGTATTTGGATG     | 842  |
| OY | 839  | ggacacactctctgaatgtgacatcotaatatctcctcgaaggtaacacacttctgtataggac | 898  |
| Db | 843  | GGCACCGAGCCCGAAGTAGCATGCACTCATTTTCTTGAAGGCGACACAGTTTCTCGTAGGAC   | 902  |
| OY | 899  | cattgcacagcgctctcttggnaatctgcaccaataactcttcctactgtcacaacactctg   | 958  |
| Db | 903  | CATCGCCAGGCTTCTCTGGAGATCTCGCCACTAATCTTCCTACACTGCTCAGACATTCCTG    | 962  |
| OY | 959  | atggacctgggaagtttactatggtttgttaactcttcctccacaaatgatggatg         | 1018 |
| Db | 963  | ATGGACCTTGGCCAGTCTCTACTGTTTGTCTATATCTCTCCACACCAATGAGGCGATG       | 1022 |
| OY | 1019 | gaagcttatgtcaaaatagacacgctgtccagagagaccccaactacgatatgaataaat     | 1078 |
| Db | 1023 | GAGGCTCAGCTCAGATATAAAGCTTCGGCGAGAGGCCCACTCGSAGGAAGAAGCTGAT       | 1082 |
| OY | 1079 | gaagaagcggaaactatgatagtacttactgattctgaatatgaltgtygtcaggtt        | 1138 |
| Db | 1083 | GAAAGA--GGAAATTTATGATGACAAATTTGTACGACTCGGACATGAGAGTGTGGCGCTC     | 1139 |
| OY | 1139 | gatgatgaacactctctctctcttcttcacaaatctgactaagttgcagaagaacatctctaa  | 1198 |
| Db | 1140 | GATGGTACACCTGTCCTCCCTTTATCCAAATCCGCTGGTGGCCAAACCAATCCGAAA        | 1199 |
| OY | 1199 | acttggtaacattacattgctgtcgtcaagaagagaccttggaactatgctcccttgaactc   | 1258 |
| Db | 1200 | ACCTGGGTGACTACTACTCTCTCGAAGAGAGAGAGACTGGAGCTACGCCGCCGGCTCCC      | 1259 |
| OY | 1259 | gcccgcgattgacagaagtataaaagtcaatatttgnaacaaatgyccttcagcggatgtgt   | 1318 |
| Db | 1260 | AGCCCCAGTGACAGAGATTATTAAGATGCTCTACTTGAACACAGTGGTCCACAGCAATTGGT   | 1319 |
| OY | 1319 | agggaagtaaaaaaagttccgatttatctgatacacaagttgaaaccttaagactcgtgaa    | 1378 |
| Db | 1320 | AGGAATATACAAAACACCTCGATTCTGCTGTACACGGATGTGAACATTTAAACATCGTAA     | 1379 |
| OY | 1379 | gcatctcaagatbaactcaggaatctctggagaccttaactctcttgsgaagttgagacaca   | 1438 |
| Db | 1380 | GCTATTCCTCGATGAATACAGGAATCTCTGGACCTTTACTTTATGGAGAAAGTTGGAGACACA  | 1439 |
| OY | 1439 | ctgttgattatatttaagaatcaagaacagacagacataataacatctaacctcacaagatc   | 1498 |
| Db | 1440 | CTTTTGTATATATTTTAAGATTAAGAGGAGACCGACCATATTAACATCTACCTCATGGAATC   | 1499 |
| OY | 1499 | actgatgtccgcgtctcttgtattatccaaggagattaccaaaagttgtaaaaactttgaagat | 1558 |
| Db | 1500 | ACTGATGTACGCGCTTTGGACCCAGGAGACCTTTAAAAGGTGGAAACATTTTAAAAGAC      | 1559 |
| OY | 1559 | tttccaattctgcagagaaatatctcaaatataatgagacagttacgtgtgaagatcgag     | 1618 |
| Db | 1560 | ATGCGCAATTTGCGCAGAGAGACACTTTCAGGTATAATGGACAGTCACTGTGGAAGATGGG    | 1619 |
| OY | 1619 | ccaactaataacagatccctcgatgctgcacccgcttatactcctagtttcgttaatalagag  | 1678 |
| Db | 1620 | CCAACCAAGTCCGATCTCTGGTCCGAGACCCGCTACTCTGAGCTCATTAATCTAGAG        | 1679 |
| OY | 1679 | agagatcttagcttcagagactcaattggccctctccctcaatcgtctacaaagaatctgtatg | 1738 |
| Db | 1680 | AAAGATCTGGCTTTCGGGACTCATTTGGCCCTCCCTCCATCTGCTCAAAAGAAATCTGTAGAC  | 1739 |
| OY | 1739 | caagaaggaacacagaataatgcacacagaaggaatgtcacctcgtttttctgtatgttat    | 1798 |
| Db | 1740 | CAAGAGGAAACCAAGTATGTCTGACACAGAAACAGTATCTCTGTTCTGTATTTGGAT        | 1799 |
| OY | 1799 | gagacacgaagctcgtgtacctacacagagaatatataacagcttctctcccaatccagcttga | 1858 |

|    |      |  |      |
|----|------|--|------|
| Db | 1800 | GAGATCAAAACCTGTGACTCTCGACAGAAATATTTACACGGCTTCTCCCTCCCAATCCGATGGA   | 1859 |
| OY | 1859 | gtcgacgtcttgaggatccagaagttccaagacctccaacatcatgcaagcaataatgcat      | 1918 |
| Db | 1860 | TTACAGCCCCCAGGATCCAGAGTTCCAAAGCTTCTTAACATATATGCACAGCATCAATGGCTAT   | 1919 |
| OY | 1919 | gtttctgatagtctgcagactggtcagttgtgttgcagaggtggacatgtaactgatactcaaac  | 1978 |
| Db | 1920 | GTMTTGTATACCTTGACGCTGTGCGTTTGTGTTCACGAGGTGGCATATCGTATCTGTA         | 1979 |
| OY | 1979 | agcaattggagcaagactgactctccctctcgtctctctctctcgtgatalaactcctcaaac    | 2038 |
| Db | 1980 | AGTGTGGAGCAGACAGAGGACTTCTCTCCGTCCTCTCTGCGCTACACCTTCAAAAC           | 2039 |
| OY | 2039 | aaatgtgtctatggaagacaacacacactatcccaattctcagggagaaactgtctctg        | 2099 |
| Db | 2040 | AAATATGGCTATGTAGAGACACACTACCTCTCTCCCTCTCAGGAGAAACGGTCTTCAATG       | 2099 |
| OY | 2099 | tcgaatggaaaaacccaagttctatgatatctggtgtgcacaaactcagaaacttctggagacaga | 2158 |
| Db | 2100 | TCATATGGAATCCACAGGTCTCTGGTCTATAGGGTGCACACACTCAGACTTGGGAAACAGA      | 2159 |
| OY | 2159 | ggcaatgacgccttaactgaaagttctcagttgtgtgacagaacaactggtgatatacag       | 2218 |
| Db | 2160 | GGGATGACAGCCTTACTGAGAGGTGTATAGTTGTGACAGGACACATTGCTGATTTATTTATAC    | 2219 |
| OY | 2219 | gacagtatggaagatatttcaagataactgtgtgagttaaaacaatgacatgtaaccaaga      | 2278 |
| Db | 2220 | AACCTTATGAGATTTTCCAGAGCTTCTTGCTGTAGTGGAAAGATGTCAATTAACCCAGA        | 2279 |
| OY | 2279 | agctctccacgaattcaagaacacctagactagcaaaagcaatttaatgcccacca           | 2338 |
| Db | 2280 | -----  | 2279 |
| OY | 2339 | ccagttctgaaacgcataacggygaataaactcgtactactcttcagtcagatacgaatag      | 2398 |
| Db | 2280 | -----GACATPAAAGCTTCTCTACTTTTTCAGCCGGAGGAGAAC                       | 2315 |
| OY | 2399 | gaattgactatgtagtatacatataagttgaatggaatgaagaagaagattttaacttat       | 2458 |
| Db | 2316 | AAATATGGCTATGTATGATTTCTTCTCAACTCGAACAACAAGAGAAATTTATACATTTTAC      | 2375 |
| OY | 2459 | gataagatatgaataatcagagcccccgaagcttccaagaagaaacacgaactatttat        | 2518 |
| Db | 2376 | GGTAGAGATGAATATCAGAACCTCCGAGCTTCAACAAGAAACCCGACATATTTCATT          | 2435 |
| OY | 2519 | gtctcagatggaagaggtctcggatattggtatgtagtagcttcccacatgltctlaagaaac    | 2578 |
| Db | 2436 | GCTCGGGTGCAGCTCTGGATTTACGGGATGACGGAATCCCCCGGGCGCTTAGAANAAC         | 2495 |
| OY | 2579 | aggagctaaagtaggcagtgctgcctcctcagttcaagaagaagtgttttccagaatttactgat  | 2638 |
| Db | 2496 | AGGGCTCAAGAACGGAGAGTGGCTCGGTTCTAAAGATGGTGTCTCCGGGAATTTCTGAC        | 2555 |
| OY | 2639 | ggctcccttactcagccctataacgltbgagaactaaatgaacatttggagactccctggg      | 2698 |
| Db | 2556 | GGCTCTTCAACGCAAGCGGTGTATCCGGGGGAGACTCAACAAACACTTGGGGCTTTGGGA       | 2615 |
| OY | 2699 | ccatataatgaagcagaagttgaagabaatalcaigttaacttccaagaatcaggctct        | 2738 |
| Db | 2616 | CCCTATCATCAAGAGCGAAGTTGAAGAACACATATGTTAACTTTAAAAACACAGCGCTCT       | 2675 |
| OY | 2739 | cgctccattctcctctatcttlaagccttatttctttagagaagaatcagaggaagaaga       | 2818 |
| Db | 2676 | CGTCCCTATTTCTTACTCTGACCTTATTTCTTATTCGGATATGATCAGGACGAAGGGCA        | 2735 |
| OY | 2819 | gaacctagaanaaaactgttgcagcctaataatgaacaacaaacttacttggaaagtgcga      | 2878 |
| Db | 2736 | GAACCTGACACAACTTCGTCCAGCCAAATGAANAACGAGACTTACTTTTGGAAAGTGCAG       | 2795 |
| OY | 2879 | catcaatggaacccaataaagaatgattgactgcaaaagccttggcttaattctctgat        | 2938 |



MOLECULE TYPE: cDNA  
US-08-717-294-42

Query Match 58.9%; Score 2603.8; DB 3; Length 4451;  
Best Local Similarity 74.8%; Pred. No. 0;  
Matches 3262; Conservative 0; Mismatches 1097; Indels 0; Gaps 0;

59 gccccagaagaatatactactggtgagcagtggaactgcatgtagggactatataatgcaagtgat 118  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db GCCACCGCCGCTACTACTGCTGGGCGCGTGGAGGCTGCTGGAGTACTATCATCAAGAGCCAC 143  
119 ctgcgtgagctgctgctgagcagcaagattctctctagagtgccaaatcttcttcattc 178  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db CTGGGGCGAGCTCCCGTGGAGCGCCGCTTCCCGCCCGCGGAGCCCAAGAGTTCCCTTC 203  
179 aacacctaagtcgtgtagcaaaaagactctgtttagaataatcaagtgtaacgttcttcaac 238  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db AACACAGCGGTGTACAGMAAAACCTGTTCGTGGAGTTCACCGACACACCTGTTCAAC 263  
239 atgcctaagcccaagcgaacctgtagtgctgtagtcttaccatccagcagtgat 298  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db ATTGCACAGCGCGCCCGCCCTGGATGGGCTGCTGGGCGCCACCATTCACAGCGCGAGTG 323  
299 taatatacagtgatcatatacacttaagaacatggtcttccactctgtaagtcctatgct 358  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db TAGACACCGGTGTGTACCTCAAGAACATGGCCAGCCACCCCGTCAAGCTGCACAGCC 383  
324 tagatagatctactctggaagcttctgagagctgtaataatgtagtcagacagtgca 418  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db GTGGGCGTACTACTGCTGAAAGCGCAGAGGCGCCGAGTACGACGACACAGACGTCCAG 443  
419 aggaagaagaagatgataagcttctccctgtagaagccataatgtagtcagagtc 478  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db CGCAGAGAGGAGGAGCAAGAGTGTCCCGGGGGAGGACCACTACGTGTGACAGTG 503  
479 ctgaagaagaatgtagcaatgtagccttgaacctgtagccttactactactatcttct 538  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db CTTAAGGAGAACGGCCCTATGTGGCCAGGACCCCTGTGCTGACCTACAGTACTGTAGC 563  
539 catgtgacctgtagtaaaagacttgaatcaggtctcatctgagccctactagtagtga 598  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db CACGTGACCTGTGAAAGATCTGAACAGCGGCTGATGCGGCTGTGCTGTGTGCG 623  
599 gaagggagcttgcgcaagaagaacacagacctgcaacaattataactacttcttgc 658  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db GAGGCGAGCTGGCCAGAGAAACCCAGACCTTGCAAGTTTATCTGCTGTTCGCC 683  
624 gattatgtagaaggaaagtgtgacctcagaacaagaacacctcctgtagtcagtagatg 718  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db GTGTTCAGACGAGGGAGAGAGCTGGCAGCGAGACTAAGAACACCTGATTCAGAGACCGC 743  
719 gatgtgcatctgctggtgacctgagcctaaatgacacagtcgaatgtagtgaacagg 778  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db GACGCCCGCAGCGCGCCGCTGGCCCAAGATGCACACCGCTTAACGGCTACGTGAACCGC 803  
779 tctctgcaagtgctgtagttagtgacagaagaatcagttatctatgtagttagttagatg 838  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db AGCTTGGCCGCGCTGATGGCTGGCCAGCGAAGCGTGTACTGTCGACGTATCGGATG 863  
839 ggcacacacctgtagagtcactcaatattctctgaaggtacacacattctgttaggaac 898  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db GGCACACCGCTGAGGTGAGCATCTTCTCTGAGAGGGCCACACCTTCTGCTGCGAC 923  
899 catgcagagcgctctgtagaatctgcacaataacttcttacttagtctcaaacacttg 958  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db CACGCGCAGGCGACCTGTGAGATCAGCCCATCTTCTGATGCTGCCAGACCTGCTG 983  
924 atggaacttgtagaagttctactgtagttagttagttagttagttagttagttagttagt 1018  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db ATTGACCTTAGGCGCATGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1043  
1019 gaagctatgtcaaaagtagcagctgtagcagaggaacccaactaagtagtaaaataat 1078

1044 GAGGCTTACGTGAAGGTGAGAGCTGCCCGCAGAGAGCCCGACGCTGCGATGAAGACAC 1103  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1079 gaagaagcggagaactatgtagttagttagttagttagttagttagttagttagttagt 1138  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1104 GAGAGGCGCGAGCTAG 1163  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1139 gatgtagcaact 1198  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1164 GACGACGACAAAGCGCCGAGCTTATCCAGATCCGACCGGTGGCCAGAAACACCTTAG 1223  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1199 acttgtagtatacatctgctgctgagaaagagagactgagactatgctctctctctct 1258  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1224 ACTTGGGTGCACTACTATGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1283  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1259 gccccagatgacagaagttataaagttcaatalttgaaacatggtccctcagcagtagt 1318  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1284 GCCCGCGACGCGCGAGCTACAGAGCCAGTACCTGAACACAGGCGCCCGACCGCATCGC 1343  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1319 aggaagtacaaaagagctgtagttagttagttagttagttagttagttagttagttagt 1378  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1344 CGCAAGTACAAAGAGTGCCTTATGCTACACCGACGAGACTTTCAAAGACCGCGAG 1403  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1379 gcatcacaatgaatcagaatcttggaccttacttcttcttcttcttcttcttcttctt 1438  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1404 GCATTCAGACGAGTCCGCGATCTCTGCGCCCGCTGCTGACGGGAGTGGCGACAC 1463  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1439 ctgtgattatataatgaatcaagaagaagcagacacataatactactactcagaatc 1498  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1464 CTGCTGATCATCTTCAAGAACACAGCCAGCGCCCTTCAACACTTACCCCGAGCATC 1523  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1499 actgattcgctctcttctgtagttagttagttagttagttagttagttagttagttagt 1558  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1524 ACCGAGTGCCTCCCTGTACAGCGCGCGCTGCGCAAGGCGTGAACACCTGAAGAGC 1583  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1559 ttccaactctgccaagaagaatataatcaataatgaagtagttagttagttagttagt 1618  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1584 TTCCCATCTGCTCCCGGAGATCTTCAAGTACAAAGTGAACCGGTGAGAGAGCGC 1643  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1619 ccaactaaatcagaatctcgtgctgtagcagcagctactactacttagttagttagttagt 1678  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1644 CCCACAAAGAGCGACCCCGCTGCTGACCGCTACCTACACACCTTGAACATGAG 1703  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1679 agagatctagcttagaagactatgtagcctctctctctctctctctctctctctct 1738  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1704 CGGACCTGCGCTCCGAGCTATGCGCCCGCTGCTGATCTCTCAAGAGAGCGTGGAC 1763  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1739 caaagagaagaacagataatgtagcagaagaagtagttagttagttagttagttagttagt 1798  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1764 CAGGCGGCGAACGATCATGAGGAGACAGAGGCAACGTGATCTGTTACGCTGTTCGAC 1823  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1799 gagaacgaagctgtagcctcacaagaataatacaagcttctctcccaatcagctgga 1858  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1824 GAGAACCGAGCTGTATCTGACCGAGAACATCCAGCGCTCTTCCGCCAACCCCGCTGC 1883  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1859 gtgagcttagaggttcaaggttccaagctcccaactctctctctctctctctctctct 1918  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1884 GTGCAGCTGGAAGATCCCGAATTCAGGCGCAACATCAATCAACACATCAACGCTTAC 1943  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1919 gttttatagttgaggttgcagttgttctgtagttagttagttagttagttagttagttagt 1978  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1944 GTGTTCAGACCGCTGAGAGTGAAGGTGTGCTGTGATGAAGTGAAGTGAAGTGAAGTGA 2003  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1979 agcatgaggaacagactgacttctctctctctctctctctctctctctctctctctct 2038  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 2004 AGCATGCGCGCGAGACGCTTCTGAGCGTGTCTTCTCCGGGTTCACCTTCAACGAC 2063  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 2039 aaaaatgtagttagaagaacacactacacctatctccatctcagaagaagtagttagttagt 2098  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 2064 AAGATGTGTAGAGAGACACCTGATCCCTGTCTTCCCGGAGAGCTGTGTTATG 2123  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 2099 tcgattgaaacccaggttcatatgtagttagttagttagttagttagttagttagttagttagt 2158  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

|    |      |   |      |
|----|------|---|------|
| Db | 2124 | TCGTATGGAGAAACCCGGGCGCTGTGGATTCTTGGGGTCCGCAACACAGCAACTTCCGAAACGCC       | 2183 |
| Oy | 2159 | ggcatagaccgccttactgaaggttcttagtltgtgcaagaacacvtgltgattatcag             | 2218 |
| Db | 2184 | GGCATACAGCCCTGTGGTAAGTCTCCAGCTGGGACAAACACGGCGACTACTGAGAG                | 2243 |
| Oy | 2219 | ggcaagtatgaagaatttccagatcttgcgtgagtaaaaaaaatgtccatttgaaaccaaga          | 2278 |
| Db | 2244 | GACAGGTACGAGGAGACATCTCCGCTACTCTCTGTCCAAAGAACAGGCATCTGACGCCCGC           | 2303 |
| Oy | 2279 | agctctcccaagatttcaagacacaccccttagcacttagagaaagcaattlaatgccacca          | 2338 |
| Db | 2304 | TGCTTCTCCCAAAACATCCGCCACCCGACGACGGGTGAGAAGCAGTTCAAGCCACCCCC             | 2363 |
| Oy | 2339 | ccagtccttgaacgcacatcaaacggggaataactcgtactactcttcagtcagatcaagaag         | 2398 |
| Db | 2364 | CCCGTGCTCAAGCGCCACCGCGCGAGATCACCGCACCCCTGCAACGACGACACAGAG               | 2423 |
| Oy | 2399 | gaaattgactatgatgatccatatacaglttgaaatgaagaagaagatttgacatttat             | 2458 |
| Db | 2424 | GAGATCGACTATCACACACACATACGCTGTGAGTGAAGGAAGGAGCTTCGACATCTTAC             | 2483 |
| Oy | 2459 | ga tggagatgaaatatagaagccccgcagcttccaagaagaacacacatatttat                | 2518 |
| Db | 2484 | GACGAGGAGAGAGAACACGAGCCCGCTCTTCCAAAGAAAACCCGCACTACTTCATC                | 2543 |
| Oy | 2519 | gctcgacgtggaagagctctvggaattatbgya tgaatgctccccaatgttctaagaac            | 2578 |
| Db | 2544 | GCGCGCGTGGAGGCGCTGTGGGACTACGGCATGAGCGACGCCGCCACGATCTGGCAC               | 2603 |
| Oy | 2579 | agggcctcagagtggaaggtctccctcagttcaagaagaattttccaaggaattacgat             | 2638 |
| Db | 2604 | CGCGCCCAAGCGCGACGCTGCCCCAGTTCAAGAGGTGTTTCAGAGATTACCGAC                  | 2663 |
| Oy | 2639 | ggctcccttactcagccctatatacgttggagaaactaaatgaacatttgggactccttggg          | 2698 |
| Db | 2664 | GCGAGTTCACCACCGACGCCCTGTACCGGGGCGAGCTGAMCAGACACCTGGGCTCTCGGC            | 2723 |
| Oy | 2699 | ccatatataagaagcagaaglttgaagaataatcagttaacttcagaatatcaagcctct            | 2758 |
| Db | 2724 | CCCTCATCTCGGCGCCGAGGGGAGGACAMCATATGTTGACTTTCGCAACCAAGCCCTCC             | 2783 |
| Oy | 2759 | cgctccattctcttcatattcagccttatattcttatagaagaatcagaagcaagsgca             | 2818 |
| Db | 2784 | CGGCGCTACTCTCTTACTCTCTCCGTGATGACTACAGGAGGACGACGCCACAGGGGCC              | 2843 |
| Oy | 2819 | gaacctagaaaaaacttctgcaagccttaatgaaaccaaacttacttcttggaaatgcga            | 2878 |
| Db | 2844 | GAGCCCGGCAAGAACTTCGTGGAAGGCCCAACGACACTTAAGACCTACTCTGGAAGTGCAG           | 2903 |
| Oy | 2879 | catcatatagcaaccaactaaagaatgatttgaactgcgaagccttvggcttatctctgat           | 2938 |
| Db | 2904 | CACCACTATGGCCCCCACAAAGGACGATTCGACTGCAAGGCTCTGGGCTACTTTCAGCGAC           | 2963 |
| Oy | 2939 | gttgcaccttgaaaaaagatbtgcactcagcgctgaattgagccctcttgcgtctcgcaact          | 2998 |
| Db | 2964 | GTGGACTTGGAGAAAGACGTGCACAGCGGCTTATGCTGCCCCCTGCTGGTGTGCACACC             | 3023 |
| Oy | 2999 | aacacactgaaacctgtcatatggagagacaagtgtacagttacaggaatttgcctgttttc          | 3058 |
| Db | 3024 | AACACCTGAACCCCCCCCCACGGGAGGACAGTGATGTCAGGAATTTGGCCGTGTTCTTC             | 3083 |
| Oy | 3059 | accatcttltgat tgaagaccaaagcctvgttacttcaactcgtaaaaatatgtgaagaagaactgcagg | 3118 |
| Db | 3084 | ACCATCTTTCGAGAGACTTAAGACTGTGTACTTTCACCGAGAACTTGGAGCCCACTGCGCG           | 3143 |
| Oy | 3119 | gtctcctgaaatatccaagatgysaagaatcccaactttaaagaagaatatgtcttccatgcga        | 3178 |
| Db | 3144 | GCCCGCTGCAACTTCAGATGAGAGATCCCACTTCAAGGAGAACTTACCGCTTCCAGGCC             | 3203 |
| Oy | 3179 | atcaatgtgtacataatgtgaataactcvtgcttgaatgtgtcaggtacaaagatc                | 3238 |
| Db | 3204 | ATCAACGGTACTATCATATGTGAGACACCTCTCGCGGCTGTGTATGGCCCAAGACACCGCATC         | 3263 |

[illegible]



|    |      |  |      |
|----|------|--|------|
| Db | 150  | CAATGAAATTAAGCTCTCCACACTGCTCTTTCTGTGCGCTTTTGCCATTCTGCTTAATGTC    | 203  |
| OY | 61   | caccagaagatactacctgggtgcagtcgaaactgtcocalggaactabataygcaaatgatct | 120  |
| Db | 210  | CACCGAAGATACACTACCTGGGTGCAGTGCAGAACTGTCATGGAGCTATATATCAAGAGATCT  | 269  |
| OY | 121  | cggtgagctgtcctgtgagcgcgaagattcctcctaaagatgyccaaaatctttccaltcaa   | 180  |
| Db | 270  | CGGTAGCTGCTGTGGAGCGCAAAATTTCCCTTAAGTCCAAAATCTTTTCATTTCAA         | 329  |
| OY | 181  | caactcaagtctgtcaaaaagaagctcgtttgttagaatcaaggttcacacttccaact      | 240  |
| Db | 330  | CACCTCACTCTGTACAAAAAGACTCTGTTGTAGAAATTCACGGTTCCACCTTTTCAACT      | 389  |
| OY | 241  | cgtlaagccaagagccacctgagatgggtctgtcgtaggctccaccatccagcgtgaggtta   | 300  |
| Db | 390  | CGCTAAGCCAAAGCCACCTGGATGGGTGCTTAGGTCCTACCATCCAGCTGAGGTTTA        | 449  |
| OY | 301  | tgaatacagttgtcatctacccttaagaacatgtgcttccactcctgtcagtcctcatgctgt  | 360  |
| Db | 450  | TGATACACTGTGTCATTCACATTAAGAACATGAGCTCCCATCTGTCAGTCCATGCTGT       | 509  |
| OY | 361  | tgtgtatccctaccacgggaagactcttgaggagagcttgaatattgaatcagaacagtcgaag | 420  |
| Db | 510  | TGGGTATCTCTACTGGAAAGCTTGTGAGGAGCTGAAATATGATGATCAACACAGTCAAG      | 569  |
| OY | 421  | ggagaaagaagatgataaagctctccctctgttggaagccatacatgctctggcagctcct    | 480  |
| Db | 570  | GGAGAAAGAAATGATTAAGTCTTCCCTGGTGGAAGCATACATATGCTGTGGCAGGTCT       | 629  |
| OY | 481  | gaaagagaatggttccaatgtgcctctgacccaatgtgcttaccactaatatcttctca      | 540  |
| Db | 630  | GAAAGAGATATGTCCAATGGCCTCTGACCACATGTCCTTACTTACTCATATCTTTTCA       | 689  |
| OY | 541  | tgtggaactgttlaaagaacttgaataggccctcatcttggaagccctactagtatgtaga    | 600  |
| Db | 690  | TGTGGACCTGGTAAABAACCTTGATTAAGCCCTCATTTGGAGCCCTACTGATGATGAGGA     | 749  |
| OY | 601  | agggaagctgcgccaaggaaaagacacagacctgtgacaaattatactactcttctgcgt     | 660  |
| Db | 750  | AGGGAGTCTGGCCCAAGGAAAAAGACACAGACTTGCACAAATTTATATCTTTTGGCT        | 809  |
| OY | 661  | attgtatgaagggaagaagtgtgcactcagaacaagaactcctctgtatgcagatagga      | 720  |
| Db | 810  | ATTGATGAAGAGGAAAGATTGGCAGCTCAGAAACAAAGAACTCCTTGATGACAGATAGGA     | 869  |
| OY | 721  | tgtcgtcatctgtcctgggcctgccttaaaatgacacacagtcgaatggttatgtaacaagtc  | 780  |
| Db | 870  | TGCTGTCACTCTCTCTGGGCTCTGGCTTAAATATGACACAGTCATATGTTATGTAAACAGTC   | 929  |
| OY | 781  | ctctccaggtctgtatgtatgcccacaggaataagctcatctgacatgtatggaatgg       | 840  |
| Db | 930  | TCTGTCCAGTGTGATTGATGGATGCCACAGAAATAGTCTATTGGCAGTGTATTTGGATGG     | 989  |
| OY | 841  | caaccactctgaagttgactcaataatctctcgaaggtccacaacttctctgtgaggaacca   | 900  |
| Db | 990  | CACACACTCTGAATGCACTCAATATTTCCTCGAAGGTCCACAAATTTCTGTGAGGAACCA     | 1049 |
| OY | 901  | tcgcgcagcggctcttgaatatctgcacaataacttccctactgctccaacactcttgat     | 960  |
| Db | 1050 | TTCGCCAGCGCTCTTGGAAATCTCGCCCAATATCTTCTTACTGCTCAAAACCTCTTAT       | 1109 |
| OY | 961  | ggaccttgaagacttctactgltttgtcatatctctccacacaacatgatatgycatgta     | 1020 |
| Db | 1110 | GGACCTTGGAGAGTTCCTACTGTTTGTGTATCTCTTCCACACAAAGATGAGGATGGA        | 1169 |
| OY | 1021 | agcttatgtcaaaagtagacagctgtccoaaggaagaccccaactagaatgaaataatgta    | 1080 |
| Db | 1170 | AGCTTATGTCAAATATACAGCTGTCCAGAGGAACCCCAACTACGAATGAAAAATTAATGA     | 1229 |
| OY | 1081 | agaagcggaagactatgatatgattcttactatcttctgaaatgagatggtgtgaagtttga   | 1140 |
| Db | 1230 | AGAAAGCGGAAGACTATGATATATATATCTTACTGATTTCTGAAATGAGATGTGTACAGTTTGA | 1289 |





|    |      |   |      |
|----|------|---|------|
| QY | 1    | catgcaaatagagctctcccaactgctctctctgtgaccttttgcaattcgtcttaagtcg     | 60   |
| Db | 150  | CATGCAAAATAGAGCTCTCCACCCTGCTTCTTCTGTGCGCTTTTGGCAATTCGCTTAGTGC     | 209  |
| QY | 61   | caaccagaatactctccctggtggtgcagttgaaactgtcatatggagactatagcaagtgacct | 120  |
| Db | 210  | CACCGAATAATCTACTCTGGGTGCAGGTGAACTGTATGGAGCTATATGCAAGAGACT         | 269  |
| QY | 121  | cggtgagctgcctgtgagacgcaagatlltccctcctagagtgccaactcttttccatlcga    | 180  |
| Db | 270  | CGGTGAGCTGCGCTGTGGAGCGCAAGATTTCCCTCTAGAGTGCCTAAATCTTTTCCATTCGA    | 329  |
| QY | 181  | caacctcagtcgtgtcaaaaaagactcgttgttagaatcaacggttccacttttcaacat      | 240  |
| Db | 330  | CACCTCAGTCGTGTCAAAAAAGACTCTGTTGTAGAAATTCAGGTTCACCTTTTCAACAT       | 389  |
| QY | 241  | gccttagaccgaagggcaaccttgataggtctgtctagttcctctaccacgaggtcgaggtta   | 300  |
| Db | 390  | CGCTTAGACCAAGGCCACCTCGATGGGTCTGCTTAGGTCCTTACCATTCAGGCTGAGGTTTA    | 449  |
| QY | 301  | tgatacagtgatcatctataccttaagaacatggtctcccatcctgtcagtcctcatcgtgt    | 360  |
| Db | 450  | TGATCAGTGTGCATTTACACTTAAGAACATGCTGCCATCCCTGTGCATCTTCATGCTGT       | 509  |
| QY | 361  | tgtgtatctactctgtgaaagctctcgaggagctgtaatatgtatgtatgcacgaactcaag    | 420  |
| Db | 510  | TGTGTATCTCTACTGTGAAAGCTTCTGTAGGAGCTGAATATGTATGTACAGCCAGCAAG       | 569  |
| QY | 421  | gagaaagaaga tga taaagctctccctggtggaagcatatcatatgtctgacagttcct     | 480  |
| Db | 570  | GAAGAAAGAAGATGATTAAGTCTTCCCTGTGTGAGAACCATATATGTCTGTGGAGTCTCT      | 629  |
| QY | 481  | gaaagaagaatggttccaatggtccttgacccaactgtgccttaccctacatatactttctca   | 540  |
| Db | 630  | GAAAGAAGATGTTCCAAAGGCTCTGACCCCACTGTGCTTACCTCATATCTTTCTCA          | 689  |
| QY | 541  | tgtgacacctgtgaaagaacttgaattcagggctcatatggagccctcactagatgttagaga   | 600  |
| Db | 690  | TGTGACACTGTGTAAGACTTGAATTCAGGCTCATTTGGAGGCTTACTAGTATGTAGAGA       | 749  |
| QY | 601  | aggagctctgagccaaagaaagacacagacacttgacacaatttactactcttttgcgt       | 660  |
| Db | 750  | AGGAGCTCTGGCCAAGAAAGACACAGACCTTGCAAAATTTATCTTTTGTCTGT             | 809  |
| QY | 661  | atttataagaaggaaagtgtggcactctcgaacaaagaaccccttgaatgcagagataggga    | 720  |
| Db | 810  | ATTTATATGAAGGAAATTTGGCACTGCAGAAACAAAGAACTCTTGAAGCAGAGATAGGA       | 869  |
| QY | 721  | tgtctcatctgctcgtggccttgacctaaatgtcacacagtcaatgttattgttaaacaggtc   | 780  |
| Db | 870  | TGCTCATCTGCTGGGGCTGGCCTTAATAATGCACACACTCAATGGTTATGTAAACAGGTCT     | 929  |
| QY | 781  | tctgcacagttctgatatgtagtcacacaggaanaatcagttcatatgtgcagtgtatggaatg  | 840  |
| Db | 930  | TCTGCACAGCTGTGATTTGATGCCACAGGAATTCAGTCTATTTGGCATGTGATTTGAATGG     | 989  |
| QY | 841  | caacactcttgaagaatggaactcaatatctctgaaaggtcacacatttctgtggaagaacca   | 900  |
| Db | 990  | CACCACTCTCTAATGTGCCTCAATATTTCTCGAAGGTACACATTTCTTTGTAGAGAACCA      | 1049 |
| QY | 901  | tcgccaagcgctcctgtgaaactctgcgccaataacttctccttactgctcacaacacttgat   | 960  |
| Db | 1050 | TCGCGACGCGTCTGTGAAATCTCGCCCAATTAATTTCTTACTCTCTCAACACACTTGTAT      | 1109 |
| QY | 961  | ggaccttggacaagtttcttactggttttgcatacatctcttccacaacaatgtagtcagga    | 1020 |
| Db | 1110 | GGACCTTGTGAAGTTTCTACGTGTTTGTGCATATCTTTCCACCAACACTGTATGTGCATGGA    | 1169 |
| QY | 1021 | agcttatgtcaaaatagacagctgtccagagaagcccaactacgaatgtaaaataatga       | 1080 |
| Db | 1170 | AGCTTATGTCAAAGTAGACACTGTCCAGAGAACCCCACTACGAATGAAAAATTAATGA        | 1229 |

```
Db 2310 CATGACCGCCTTACTAGAGGTTTCTAGTTGTGACAGACACTGSGATTTATTACGAGCA 2369
Qy 2221 cagttatgaagataattcagcatacttgctgagtaaaaaacaatgcatltgaaccaagaag 2280
Db 2370 CAGTTATGAAGATATTTTCAGCATACTTGTGATAAAAAACAATGCCATTGAACCAAGAAg 2429
Qy 2281 ctctcccaagaattcaagacacccctagcaactaggcaaaagcaattlaatgccaccga 2338
Db 2430 CTTCTCCCAAGATTCAAGACACCCCTAGCAGCTAGGCAAAAGCAATTTAATGCCACCACA 2487
```

Search completed: November 18, 2001, 04:58:30  
Job time: 35512 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2001, 05:05:42 ; Search time 919.72 Seconds  
(without alignments)  
3015.528 Million cell updates/sec

Title: US-09-689-430-1\_COPY\_419\_4835  
Perfect score: 4417  
Sequence: 1 catgcaatagagctctcca.....agccacagacctctactgta 4417

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: N\_Geneseq\_0601.\*  
2: /SIDS8/gcgdata/geneseq/geneseqn/NA1980.DAT.\*  
3: /SIDS8/gcgdata/geneseq/geneseqn/NA1981.DAT.\*  
4: /SIDS8/gcgdata/geneseq/geneseqn/NA1982.DAT.\*  
5: /SIDS8/gcgdata/geneseq/geneseqn/NA1983.DAT.\*  
6: /SIDS8/gcgdata/geneseq/geneseqn/NA1984.DAT.\*  
7: /SIDS8/gcgdata/geneseq/geneseqn/NA1985.DAT.\*  
8: /SIDS8/gcgdata/geneseq/geneseqn/NA1986.DAT.\*  
9: /SIDS8/gcgdata/geneseq/geneseqn/NA1987.DAT.\*  
10: /SIDS8/gcgdata/geneseq/geneseqn/NA1988.DAT.\*  
11: /SIDS8/gcgdata/geneseq/geneseqn/NA1989.DAT.\*  
12: /SIDS8/gcgdata/geneseq/geneseqn/NA1990.DAT.\*  
13: /SIDS8/gcgdata/geneseq/geneseqn/NA1991.DAT.\*  
14: /SIDS8/gcgdata/geneseq/geneseqn/NA1992.DAT.\*  
15: /SIDS8/gcgdata/geneseq/geneseqn/NA1993.DAT.\*  
16: /SIDS8/gcgdata/geneseq/geneseqn/NA1994.DAT.\*  
17: /SIDS8/gcgdata/geneseq/geneseqn/NA1995.DAT.\*  
18: /SIDS8/gcgdata/geneseq/geneseqn/NA1996.DAT.\*  
19: /SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT.\*  
20: /SIDS8/gcgdata/geneseq/geneseqn/NA1998.DAT.\*  
21: /SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT.\*  
22: /SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT.\*  
23: /SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1          | 4412.8 | 99.9        | 4629   | 20    | AAx88293    |
| 2          | 4407.4 | 99.8        | 9354   | 18    | AA773164    |
| 3          | 4355.2 | 98.6        | 4629   | 16    | AA076016    |
| 4          | 4351   | 98.5        | 4670   | 19    | AAV23339    |
| 5          | 4321.4 | 97.8        | 4832   | 19    | AAV19581    |
| 6          | 4321.4 | 97.8        | 4832   | 19    | AAV15338    |
| 7          | 4321.4 | 97.8        | 9164   | 20    | AAx82259    |
| 8          | 4319.8 | 97.8        | 4999   | 21    | AA000122    |
| 9          | 4315   | 97.7        | 12445  | 21    | AA49232     |
| 10         | 4249.4 | 96.2        | 11933  | 21    | AA000121    |
| 11         | 4173   | 94.5        | 4275   | 10    | AA990654    |

|    |        |      |       |    |          |
|----|--------|------|-------|----|----------|
| 12 | 4171.6 | 94.4 | 4275  | 9  | AA80446  |
| 13 | 4170   | 94.4 | 4272  | 9  | AA80447  |
| 14 | 4161   | 94.2 | 4373  | 20 | AA82258  |
| 15 | 4152   | 94.0 | 4830  | 9  | AA81544  |
| 16 | 4108.8 | 93.0 | 4545  | 9  | AA80444  |
| 17 | 3743.8 | 84.8 | 5035  | 18 | AA69811  |
| 18 | 3708.4 | 84.0 | 4616  | 9  | AA81545  |
| 19 | 3678   | 83.3 | 5094  | 21 | AA49231  |
| 20 | 3189.8 | 72.2 | 4334  | 19 | AAV12142 |
| 21 | 3189.8 | 72.2 | 4334  | 20 | AAV91195 |
| 22 | 3189.8 | 72.2 | 4334  | 22 | AA690541 |
| 23 | 2603.8 | 58.9 | 4451  | 19 | AAV23288 |
| 24 | 2336.4 | 52.9 | 9009  | 14 | AA050185 |
| 25 | 2336.4 | 52.9 | 9009  | 18 | AA61548  |
| 26 | 2336.4 | 52.9 | 9009  | 19 | AAV25810 |
| 27 | 2336.4 | 52.9 | 9009  | 20 | AAV91162 |
| 28 | 2336.4 | 52.9 | 9009  | 22 | AA690508 |
| 29 | 2334.8 | 52.9 | 6300  | 17 | AA603571 |
| 30 | 2334.8 | 52.9 | 8241  | 9  | AA81439  |
| 31 | 2334.8 | 52.9 | 8241  | 9  | AA81096  |
| 32 | 2334.8 | 52.9 | 8975  | 6  | AA80054  |
| 33 | 2334.8 | 52.9 | 8975  | 21 | AA286604 |
| 34 | 2334.8 | 52.9 | 9029  | 22 | AA60309  |
| 35 | 2334.8 | 52.9 | 9068  | 19 | AAV15359 |
| 36 | 2334.8 | 52.9 | 9080  | 19 | AAV19580 |
| 37 | 2334.8 | 52.9 | 11846 | 20 | AAx82261 |
| 38 | 2333.8 | 52.8 | 7053  | 18 | AA651357 |
| 39 | 2333.2 | 52.8 | 8967  | 17 | AA631031 |
| 40 | 2333.2 | 52.8 | 8967  | 22 | AA687526 |
| 41 | 2332.2 | 52.8 | 9009  | 19 | AAV18884 |
| 42 | 2332.2 | 52.6 | 7056  | 15 | AA066615 |
| 43 | 2309.8 | 52.3 | 7440  | 7  | AA60689  |
| 44 | 2309.8 | 52.3 | 12022 | 20 | AAx82260 |
| 45 | 2272.4 | 51.4 | 7440  | 6  | AA850375 |

#### ALIGNMENTS

|          |  |
|----------|--|
| RESULT   | 1  |
| AAx88293 | standard; DNA: 4629 BP.  |
| XX       |  |
| AC       | AAx88293:  |
| XX       |  |
| DT       | 24-SEP-1999 (first entry)  |
| XX       |  |
| DE       | Human Factor VIII with B-domain deleted cDNA.                          |
| XX       |  |
| KW       | Adenoviral vector: Factor VIII; Factor IX; clotting factor; treatment; |
| KW       | haemostatic; hemophilia A; haemophilia B; gene therapy; ss.            |
| OS       | Homo sapiens.  |
| XX       |  |
| PN       | US5935935-A.   |
| XX       |  |
| PD       | 10-AUG-1999.   |
| XX       |  |
| PF       | 07-JUN-1995; 9505-0484891.   |
| XX       |  |
| PR       | 07-JUN-1995; 9505-0484891.   |
| PR       | 10-JUN-1993; 9305-0074920.   |
| PR       | 25-MAR-1994; 9405-0218335.   |
| PA       | (GENE-) GENETIC THERAPY INC.   |
| PI       | Connolly S, Kaleo M, Smith T;  |
| XX       |  |
| DR       | WPI; 1999-457617/38.   |
| XX       |  |
| PT       | Adenoviral vectors useful for treating hemophilia                      |
| XX       |  |
| PS       | Example 1; Column 53-58; 90pp; English.                                |

XX This invention describes novel adenoviral vectors comprising at least  
CC one DNA sequence encoding a clotting factor (Factor IX or Factor VIII).  
CC The vectors of the invention have haemostatic activity. The vectors are  
CC useful for the treatment of hemophilia A or hemophilia B by gene therapy.  
CC This sequence represents human Factor VIII cDNA which has the B domain  
CC deleted.  
XX  
SQ Sequence 4629 BP; 1319 A; 1050 C; 1015 G; 1245 T; 0 other:  
  
Query Match 99.9%; Score 4412.8; DB 20; Length 4629;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 2 atgcaaatagagctctccacgtctctctctctgctcttcttgagatctgcttagtgc 61  
DB 1 atgcaaatagagctctccacgtctctctctctgctcttcttgagatctgcttagtgc 60  
  
QY 62 accgaagatactactctggtgacgtggaactgcatgtggaactatgcaagtgatctc 121  
DB 61 acaagaagatactactctggtgacgtggaactgcatgtggaactatgcaagtgatctc 120  
  
QY 122 ggtgagctgctgtggaacgaagatctccctctagatggtgcaaatcttccatcaac 181  
DB 121 ggtgagctgctgtggaacgaagatctccctctagatggtgcaaatcttccatcaac 180  
  
QY 182 acccagctcggtgtaaaaagaactctgtttgtagaattcaggttccacgttccaac 241  
DB 181 acccagctcggtgtaaaaagaactctgtttgtagaattcaggttccacgttccaac 240  
  
QY 242 gctaagcaaggtccaccctggaatggtctgtaaggtccacatccaggctgagttat 301  
DB 241 gctaagcaaggtccaccctggaatggtctgtaaggtccacatccaggctgagttat 300  
  
QY 302 gatcaatgtgcatctacaccttaagaacatggtcttccatctctgtcagtcttaatgctgt 361  
DB 301 gatcaatgtgcatctacaccttaagaacatggtcttccatctctgtcagtcttaatgctgt 360  
  
QY 362 ggtatccctactgtaaaagctctctgaggagactgataatgatacagacagccaag 421  
DB 361 ggtatccctactgtaaaagctctctgaggagactgataatgatacagacagccaag 420  
  
QY 422 gagaaagaagaatgaataagctctccctggtggaagccaataatgctgaggtccgt 481  
DB 421 gagaaagaagaatgaataagctctccctggtggaagccaataatgctgaggtccgt 480  
  
QY 482 aaagaagaatggtccaatggtccttgacccactgtgcttactactatatacttctcat 541  
DB 481 aaagaagaatggtccaatggtccttgacccactgtgcttactactatatacttctcat 540  
  
QY 542 gtgacactgtgtaaaagactgaattcaaggtcctcatgtgagccctactagatgtaga 601  
DB 541 gtgacactgtgtaaaagactgaattcaaggtcctcatgtgagccctactagatgtaga 600  
  
QY 602 gggagctctggccaaggaagaacacagacactgtgcaaatatatactacttcttgctga 661  
DB 601 gggagctctggccaaggaagaacacagacactgtgcaaatatatactacttcttgctga 660  
  
QY 662 ttgtatgaagggaagaatgtgcaactcagaacaaagaactcctgtatgcaagataggat 721  
DB 661 ttgtatgaagggaagaatgtgcaactcagaacaaagaactcctgtatgcaagataggat 720  
  
QY 722 gctgcatctgctgggctggccttaaaatgcaacagtcacatgtgtatgtaaaaggtct 781  
DB 721 gctgcatctgctgggctggccttaaaatgcaacagtcacatgtgtatgtaaaaggtct 780  
  
QY 782 ctgccaagctctgatatgtgacacaggaatcagtcatactggcatgtatgtgaattggc 841  
DB 781 ctgccaagctctgatatgtgacacaggaatcagtcatactggcatgtatgtgaattggc 840  
  
QY 842 accactctgaagtgacactcaatatctctgaaagttacaacttctgtgaggaacat 901  
DB 842 accactctgaagtgacactcaatatctctgaaagttacaacttctgtgaggaacat 901

DB 841 accactctgaagtgacactcaatatctctgaaagttacaacatcttctgtgaggaacat 900  
QY 902 cgcagagctcctcttggaatctcgccaataacttctactactgctcaaacctttagt 961  
DB 901 cgcagagctcctcttggaatctcgccaataacttctactactgctcaaacctttagt 960  
  
QY 962 gaccttgaagttctactgtttgtcatatctcttccccaacatgatagtgagaa 1021  
DB 961 gaccttgaagttctactgtttgtcatatctcttccccaacatgatagtgagaa 1020  
  
QY 1022 gcttatgtcaaatgatacagctgtctcagaagaccccaactacgaatgaanaaatga 1081  
DB 1021 gcttatgtcaaatgatacagctgtctcagaagaccccaactacgaatgaanaaatga 1080  
  
QY 1082 gaagcggagacataatgataatcttactgtatcttgaaatggatgtggtcaggtttag 1141  
DB 1081 gaagcggagacataatgataatcttactgtatcttgaaatggatgtggtcaggtttag 1140  
  
QY 1142 gatgacaactcctctctcttataccaatctgctcagtttgccaagaagcaltcctaaact 1201  
DB 1141 gatgacaactcctctctcttataccaatctgctcagtttgccaagaagcaltcctaaact 1200  
  
QY 1202 tgggtacattacatgtgctgctggaagaggaactgggaactatgtctccttagtctggc 1261  
DB 1201 tgggtacattacatgtgctgctggaagaggaactgggaactatgtctccttagtctggc 1260  
  
QY 1262 ccggaatgacagaagttataaagaatcaaatgtgacaatgctccctcagcgatgtgtag 1321  
DB 1261 ccggaatgacagaagttataaagaatcaaatgtgacaatgctccctcagcgatgtgtag 1320  
  
QY 1322 aagtaaaaaaagtcogattatgtacacacagatgaacactttaaagctgtgaaact 1381  
DB 1321 aagtaaaaaaagtcogattatgtacacacagatgaacactttaaagctgtgaaact 1380  
  
QY 1382 attcggcatgaatcoggaatcttggaccttacttataatggggaagtgtgagacaactg 1441  
DB 1381 attcggcatgaatcoggaatcttggaccttacttataatggggaagtgtgagacaactg 1440  
  
QY 1442 ttgatatatttaagaatacgaagcagacacataacactacccctccogataact 1501  
DB 1441 ttgatatatttaagaatacgaagcagacacataacactacccctccogataact 1500  
  
QY 1502 gatgtcgtcctctgtatccaagagatcccaaaaggtgtaaaacatttgaagattt 1561  
DB 1501 gatgtcgtcctctgtatccaagagatcccaaaaggtgtaaaacatttgaagattt 1560  
  
QY 1562 ccaattctgccaaggaataatccaataatggaacagatgtagaagatgggcca 1621  
DB 1561 ccaattctgccaaggaataatccaataatggaacagatgtagaagatgggcca 1620  
  
QY 1622 actaatcagatccctcggtgctgacccgctactacttaatttcgttaatatgaga 1681  
DB 1621 actaatcagatccctcggtgctgacccgctactacttaatttcgttaatatgaga 1680  
  
QY 1682 gatctagcttaaggaatcatgtgcccctcctcatctgtctacaagaatctgtatgata 1741  
DB 1681 gatctagcttaaggaatcatgtgcccctcctcatctgtctacaagaatctgtatgata 1740  
  
QY 1742 agaggaacacagataatgtacagacagaaggaatgtaactccgtttctgtatttgtag 1801  
DB 1741 agaggaacacagataatgtacagacagaaggaatgtaactccgtttctgtatttgtag 1800  
  
QY 1802 aaccgaagctgtactacctacagagaataataacagcttctccccaatccagctgtag 1861  
DB 1801 aaccgaagctgtactacctacagagaataataacagcttctccccaatccagctgtag 1860  
  
QY 1862 cagcttgaagttccagagttccaaagctccaaatcatalgacagatcaatgtctatgtc 1921  
DB 1861 cagcttgaagttccagagttccaaagctccaaatcatalgacagatcaatgtctatgtc 1920  
  
QY 1922 tttagatgttgcaggtgtgagttgtttgtgataaggttgagatactgtacatttcaagc 1981  
DB 1921 tttagatgttgcaggtgtgagttgtttgtgataaggttgagatactgtacatttcaagc 1980

|    |      |   |      |
|----|------|---|------|
| OY | 1982 | atvgagacagactgactcttccttcgtctctcgtctcgtatatacctcaaacaaa               | 2041 |
| Db | 1981 | atvgagacacagactgactcttccttcgtctcttcctcgtatatacctcaaacaaa              | 2040 |
| OY | 2042 | atggtctatbaaagacaaacacccattccattctcgaaggaaacgttttatgctg               | 2101 |
| Db | 2041 | atggtctatbaaagacaaacacccattccattctcgaaggaaacgtttctatgctg              | 2100 |
| OY | 2102 | atvgaaaaaccagagctctatgfatctctgvggtgycacaaactcgaactctcgaaacagagc       | 2161 |
| Db | 2101 | atvgaaaaaccagagctctatgfatctctgvggtgycacaaactcgaactctcgaaacagagc       | 2160 |
| OY | 2162 | atgagccgccttactgaaaggtttcttagtttgtgtgacgaagaacactggtatatacgaagac      | 2221 |
| Db | 2161 | atgagccgccttactgaaaggtttcttagtttgtgtgacgaagaacactggtatatacgaagac      | 2220 |
| OY | 2222 | agttatbaaagatalttcgaacatactgtctgtagtaaaacaatgycattgaaacgaagagc        | 2281 |
| Db | 2221 | agttatbaaagatalttcgaacatactgtctgtagtaaaacaatgycattgaaacgaagagc        | 2280 |
| OY | 2282 | ttctcccaagaattccaagacaccccttagcactaagycaaaagcaattaatgccaacccaca       | 2341 |
| Db | 2281 | ttctcccaagaattccaagacaccccttagcactaagycaaaagcaattaatgccaacccaca       | 2340 |
| OY | 2342 | gtctctgaaacgcacatccaagcggaaactctgtactactcttcagctcgaatcaagaagaa        | 2401 |
| Db | 2341 | gtctctgaaacgcacatccaagcggaaactctgtactactcttcagctcgaatcaagaagaa        | 2400 |
| OY | 2402 | atgactatgatataccaatatacagttctgtaaatgaaagaaagaaatttggacattatgat        | 2461 |
| Db | 2401 | atgactatgatataccaatatacagttctgtaaatgaaagaaagaaatttggacattatgat        | 2460 |
| OY | 2462 | gaggaataaataataagagcccccgacgtttcaaaagaaacacgcacatatattatgtct          | 2521 |
| Db | 2461 | gaggaataaataataagagcccccgacgtttcaaaagaaacacgcacatatattatgtct          | 2520 |
| OY | 2522 | gcagctggagagagctctctgvggattatgvggagtgaatgaactccccaatgctcttaagaacag    | 2581 |
| Db | 2521 | gcagctggagagagctctctgvggattatgvggagtgaatgaactccccaatgctcttaagaacag    | 2580 |
| OY | 2582 | gctcaagagctgcaagctgtccctcagttccaagaagatgtgtttcccaagaaattactgatatgc    | 2641 |
| Db | 2581 | gctcaagagctgcaagctgtccctcagttccaagaagatgtgtttcccaagaaattactgatatgc    | 2640 |
| OY | 2642 | tcctcttctcagccctataacgctggaagacctaataatgaaacattgvgactctcctgvgggca     | 2701 |
| Db | 2641 | tcctcttctcagccctataacgctggaagacctaataatgaaacattgvgactctcctgvgggca     | 2700 |
| OY | 2702 | tataataagcagaaagtgtgaagataataatcatggttaactttcagaatactcagagctctcgt     | 2761 |
| Db | 2701 | tataataagcagaaagtgtgaagataataatcatggttaactttcagaatactcagagctctcgt     | 2760 |
| OY | 2762 | ccctattccctctcattctcagctccttattttctttatgvggaaatcaagaagcagaagcgaa      | 2821 |
| Db | 2761 | ccctattccctctcattctcagctccttattttctttatgvggaaatcaagaagcagaagcgaa      | 2820 |
| OY | 2822 | cctagaaaaaacttgttccaagccttaatgtaaaacaaaacttacttttggaaagtccaacat       | 2881 |
| Db | 2821 | cctagaaaaaacttgttccaagccttaatgtaaaacaaaacttacttttggaaagtccaacat       | 2880 |
| OY | 2882 | catatggcaccaccataaagatgaagtttgaactgtcaaaagcttvggcttatattctcgtatgct    | 2941 |
| Db | 2881 | catatggcaccaccataaagatgaagtttgaactgtcaaaagcttvggcttatattctcgtatgct    | 2940 |
| OY | 2942 | gacctggaaaaaagatgctgcaactcagaagcctggatctggaaacccctctcgtctgccaactaac   | 3001 |
| Db | 2941 | gacctggaaaaaagatgctgcaactcagaagcctggatctggaaacccctctcgtctgccaactaac   | 3000 |
| OY | 3002 | acactgtaacccctgctcatgvggagagaagaatgacagtacagaagaattgctcgtgttttccac    | 3061 |
| Db | 3001 | acactgtaacccctgctcatgvggagagaagaagaatgacagtacagaagaattgctcgtgttttccac | 3060 |

|    |      |  |      |
|----|------|--|------|
| QY | 3062 | atcttttgatgagccaaaagctggttacttcaactgaaataatggaagaacatgcagggct      | 3121 |
| DB | 3061 | atcttttgatgagccaaaagctggttacttcaactgaaataatggaagaacatgcagggct      | 3120 |
| QY | 3122 | ccctgcgaatcacagatggaagatcccaacttttaagagaattatgcgttcattgcatac       | 3188 |
| DB | 3121 | ccctgcgaatcacagatggaagatcccaacttttaagagaattatgcgttcattgcatac       | 3180 |
| QY | 3182 | aatgctacaatgagatgatacctacccgcttgtagtaatggtctcaagatccaagatctga      | 3241 |
| DB | 3181 | aatgctacaatgagatgatacctacccgcttgtagtaatggtctcaagatccaagatctga      | 3240 |
| QY | 3242 | tgtatactgctcaagatggtgcagcaatgaaacaccatcttatcaatttcasgtggaat        | 3301 |
| DB | 3241 | tgtatactgctcaagatggtgcagcaatgaaacaccatcttatcaatttcasgtggaat        | 3300 |
| QY | 3302 | gtgttcacatgtaagaaaaaaggagatataaaatgacactgtaacatatctatccacgt        | 3361 |
| DB | 3301 | gtgttcacatgtaagaaaaaaggagatataaaatgacactgtaacatatctatccacgt        | 3360 |
| QY | 3362 | gtttcttgagacaagtggaaatgtatccctcccaagcctggaaatttggcgggttggaatgcct   | 3422 |
| DB | 3361 | gtttcttgagacaagtggaaatgtatccctcccaagcctggaaatttggcgggttggaatgcct   | 3420 |
| QY | 3422 | attggcgaagcatctacatgctgtggaatgagcacactttctcgtgtgtaagacaataagt      | 3481 |
| DB | 3421 | attggcgaagcatctacatgctgtggaatgagcacactttctcgtgtgtaagacaataagt      | 3480 |
| QY | 3482 | cagatccccctgggaatggtctctcgagcaaatatagatatttcaagttaacagcttcaga      | 3541 |
| DB | 3481 | cagatccccctgggaatggtctctcgagcaaatatagatatttcaagttaacagcttcaga      | 3540 |
| QY | 3542 | caataatgacagtggtgcccccaagcctggcagagcttcattatccgagatccaataatgc      | 3601 |
| DB | 3541 | caataatgacagtggtgcccccaagcctggcagagcttcattatccgagatccaataatgc      | 3600 |
| QY | 3602 | tggagcacacaaggagcccttttcttgatataaaggatgatactgtttggcaccaaatgata     | 3661 |
| DB | 3601 | tggagcacacaaggagcccttttcttgatataaaggatgatactgtttggcaccaaatgata     | 3660 |
| QY | 3662 | caagcgcatacaagaccagaagtgcccgtaagaattctccaagctctacatctctaagttt      | 3721 |
| DB | 3661 | caagcgcatacaagaccagaagtgcccgtaagaattctccaagctctacatctctaagttt      | 3720 |
| QY | 3722 | atcaatcatgtaatgctcttgatggagaagaatggcagactatccaggaatccactctga       | 3781 |
| DB | 3721 | atcaatcatgtaatgctcttgatggagaagaatggcagactatccaggaatccactctga       | 3780 |
| QY | 3782 | accttaaatggtctctcttggcgaatgtgatactactgagataaaacaataatttttaac       | 3841 |
| DB | 3781 | accttaaatggtctctcttggcgaatgtgatactactgagataaaacaataatttttaac       | 3840 |
| QY | 3842 | cttccaattatgtctgataatactccgttttgaccacaacctatataagatttcagagact      | 3901 |
| DB | 3841 | cttccaattatgtctgataatactccgttttgaccacaacctatataagatttcagagact      | 3900 |
| QY | 3902 | cttcgcataggagtgtaatgggctgtgatttaaatagtttgcaagcatgcgcatctgggaatgag  | 3961 |
| DB | 3901 | cttcgcataggagtgtaatgggctgtgatttaaatagtttgcaagcatgcgcatctgggaatgag  | 3960 |
| QY | 3962 | agtaaaagaatatatagatgtagcagatthactgcttcaatccactttacaaatagtlttgc     | 4021 |
| DB | 3961 | agtaaaagaatatatagatgtagcagatthactgcttcaatccactttacaaatagtlttgc     | 4020 |
| QY | 4022 | acctggtctccttcaaaaagctcgactcaccttccaaagggaagttaatgctctgagacct      | 4081 |
| DB | 4021 | acctggtctccttcaaaaagctcgactcaccttccaaagggaagttaatgctctgagacct      | 4080 |
| QY | 4082 | caggttgataataatccaaaagggtggctgtagcaagtgaacttccagaagaacaatggaagtcca | 4141 |
| DB | 4081 | caggttgataataatccaaaagggtggctgtagcaagtgaacttccagaagaacaatggaagtcca | 4140 |
| QY | 4142 | ggagtaacttactcagggaagtaaaactctgtcttaccagaatgtlaatgtgaagagttccct    | 4201 |

```

Db 4141 ggaagtaactacgagggagtaaaatctgtctaccagcatgtaaggaagatccctc 4200
Oy 4202 atcccaagcagtcgaagatgagccatcagtcgagactctctttttcagaatgcaagtaag 4261
Db 4201 atcccaagcagtcgaagatgagccatcagtcgagactctctttttcagaatgcaagtaag 4260
Oy 4262 gttttcagggaaatcaagactccttcacacctgtgtgtgaactctcttagaccacgcta 4321
Db 4261 gttttcagggaaatcaagactccttcacacctgtgtgtgaactctcttagaccacgcta 4320
Oy 4322 ctgactcgtcctccttcgaattcaccgccagagttgtgtgacacaattgacctgagatg 4381
Db 4321 ctgactcgtcctccttcgaattcaccgccagagttgtgtgacacaattgacctgagatg 4380
Oy 4382 gaggtctgggctcgagagcaagcactctactga 4417
Db 4381 gaggtctgggctcgagagcaagcactctactga 4416

RESULT 2
AA73164
ID AA73164 standard; CDNA: 9354 BP.
AC AA73164;
XX
XX 08-APR-1998 (first entry)
XX
DE cDNA encoding human B-domain deleted factor VIII.
XX
XX Post-translational regulatory element; PRE; enhancer II; intronless gene;
KW surface antigen gene; cytoplasmic accumulation; targeted delivery;
KW near consensus splice sequence; blood coagulation factor; factor VIII;
XX factor IX; ss.
XX
OS Homo sapiens.
XX
XX Location/Qualifiers
FH Key 2965..7380
FH CDS
FT misc_feature
FT /tag= a
FT /note= "5' near consensus site"
FT /tag= b
FT /note= "5' near consensus site"
FT /tag= c
FT /note= "5' near consensus site"
FT /tag= d
FT /note= "5' near consensus site"
FT /tag= e
FT /note= "5' near consensus site"
FT /tag= f
FT /note= "5' near consensus site"
FT /tag= g
FT /note= "5' near consensus site"
FT /tag= h
FT /note= "3' near consensus site"
FT /tag= i
FT /note= "3' near consensus site"
FT /tag= j
FT /note= "3' near consensus site"
FT /tag= k
FT /note= "3' near consensus site"
FT /tag= l
FT /note= "3' near consensus site"
FT /tag= m
FT /note= "3' near consensus site"
FT /tag= n
FT /note= "3' near consensus site"
FT /tag= o
FT /note= "3' near consensus site"
FT /tag= p
FT /note= "3' near consensus site"
FT /tag= q
FT /note= "3' near consensus site"
FT /tag= r
FT /note= "3' near consensus site"
FT /tag= s
FT /note= "3' near consensus site"
FT /tag= t
FT /note= "3' near consensus site"
FT /tag= u
FT /note= "PRE sequence"

```

```

FT misc_feature 5604..5632
FT /tag= m
FT /note= "3' near consensus site"
FT misc_feature 5717..5745
FT /tag= n
FT /note= "3' near consensus site"
FT misc_feature 6239..6258
FT /tag= o
FT /note= "3' near consensus site"
FT misc_feature 6558..6682
FT /tag= p
FT /note= "3' near consensus site"
FT misc_feature 7159..7176
FT /tag= q
FT /note= "3' near consensus site"
FT misc_feature 7196..7209
FT /tag= r
FT /note= "3' near consensus site"
FT misc_feature 7289..7315
FT /tag= s
FT /note= "3' near consensus site"
FT misc_feature 7411..7429
FT /tag= t
FT /note= "3' near consensus site"
FT misc_feature 7611..8197
FT /tag= u
FT /note= "PRE sequence"

W09733994-A1.
18-SEP-1997.
10-MAR-1997; 97WO-US03561.
11-MAR-1996; 96US-0683839.
(IMMU-) IMMUNE RESPONSE CORP.
Bidlinauer S, ILL CR.
WPI. 1997-470874/43.
P-PSDB; AAW23414.

Vector for increased expression of intronless genes - comprises
intronless gene with at least one near consensus splice sequence, a
promoter and at least one viral cis-acting post-transcriptional
regulatory element

Example 1: Pages 21-31; 59pp; English.

The present sequence represents human B-domain deleted factor VIII
cDNA, and a post-translational regulatory element (PRE) of the
Hepatitis B virus, which is present 3' of the stop codon for factor VIII.
PRE sequences have been shown to function in cis to increase the
steady-state levels of surface gene transcripts by facilitating
cytoplasmic accumulation of these transcripts. The present sequence
is part of a novel vector, comprising an intronless gene containing
1 or more near consensus splice sequences operably linked to a
promoter sequence so that the gene is transcribed in a cell.
Intronless gene transcripts which contain near consensus splice site
sequences are believed to get tied up in the nucleus of the cell where
splicing occurs, rather than being transported to the cytoplasm where
they can be translated into proteins. The PRE sequences are transcribed
along with the gene, causing export of the gene transcript from the
nucleus into the cytoplasm of the cell. The vector can be used
to increase the expression of an intronless gene containing at least one
near consensus splice sites, preferably cDNA encoding a blood coagulation
factor, particularly factor VIII or IX. The complex allows the targeted
delivery of the vector to a specific cell, e.g. hepatocytes when the
ligand is an asialoglycoprotein which binds the asialoglycoprotein
receptor present on their surface.

Sequence 9354 BP; 2506 A; 2239 C; 2161 G; 2448 T; 0 other;

```



|                            |        |               |           |              |
|----------------------------|--------|---------------|-----------|--------------|
| Query Match                | 99.8%; | Score 4407.4; | DB 18;    | Length 9354; |
| Best Local Similarity      | 99.9%; | Pred. No. 0;  |           |              |
| Matches 4411; Conservative | 0;     | Mismatches 6; | Indels 0; | Gaps 0;      |

|    |      |  |      |
|----|------|--|------|
| QY | 1    | catgcaaatagaagctctccacccgcgtcttcctctgagcattctgcttaagtc                 | 60   |
| Db | 2964 | caatggtttatgagctctccacccgcgtcttcctctgagcattctgcttaagtc                 | 3023 |
| QY | 61   | caccagaaga taactacctggtgag tggaaactgcatatgagcattatgcagaagtgatct        | 120  |
| Db | 3024 | caccagaaga taactacctggtgag tggaaactgcatatgagcattatgcagaagtgatct        | 3083 |
| QY | 121  | cggatgagctgctgtgtgacgcaagaattctctctctagatgtgccaaaculttccatcaa          | 180  |
| Db | 3084 | cggatgagctgctgtgtgacgcaagaattctctctctagatgtgccaaaculttccatcaa          | 3143 |
| QY | 181  | cacctatgctgtgtataaaaaagactctgtttgttaaatccaaggttgaaccttttccaat          | 240  |
| Db | 3144 | cacctatgctgtgtataaaaaagactctgtttgttaaatccaaggttgaaccttttccaat          | 3203 |
| QY | 241  | cgtataagccaaagccacccctgtagtggctgtctgtaagctcttaaccatccagcttgagttta      | 300  |
| Db | 3204 | cgtataagccaaagccacccctgtagtggctgtctgtaagctcttaaccatccagcttgagttta      | 3263 |
| QY | 301  | tgatacagtggtcattatacacttaagaacaatggcttcccatccctgtcagttctcaatgctgt      | 360  |
| Db | 3264 | tgatacagtggtcattatacacttaagaacaatggcttcccatccctgtcagttctcaatgctgt      | 3323 |
| QY | 361  | tgtgtgtatcttctacttgnaaaagctctctgaaggagctgaatatgtgtgttaagaccagtcagaag   | 420  |
| Db | 3324 | tgtgtgtatcttctacttgnaaaagctctctgaaggagctgaatatgtgtgttaagaccagtcagaag   | 3383 |
| QY | 421  | ggagaaagaagaatgataaagctctctccctgtgtgaaagccataataatgtctgcagcttct        | 480  |
| Db | 3384 | ggagaaagaagaatgataaagctctctccctgtgtgaaagccataataatgtctgcagcttct        | 3443 |
| QY | 481  | gaaagaggaataggtcccaatgagctctctgaccccatgtgcttactactcatatctttctca        | 540  |
| Db | 3444 | gaaagaggaataggtcccaatgagctctctgaccccatgtgcttactactcatatctttctca        | 3503 |
| QY | 541  | tgtgtgacctgtgtataaagaacttgaaatccaagctctcatgtgagccctacagatgtataaga      | 600  |
| Db | 3504 | tgtgtgacctgtgtataaagaacttgaaatccaagctctcatgtgagccctacagatgtataaga      | 3563 |
| QY | 601  | aggagagctctgcccagaagaaagaacacagaacctctgcacaaattatatactacttttgcgt       | 660  |
| Db | 3564 | aggagagctctgcccagaagaaagaacacagaacctctgcacaaattatatactacttttgcgt       | 3623 |
| QY | 661  | atttgatggaagggaagaagtgtgagctccagaacaaagaactcttgatgcacgagataagga        | 720  |
| Db | 3624 | atttgatggaagggaagaagtgtgagctccagaacaaagaactcttgatgcacgagataagga        | 3683 |
| QY | 721  | tgtctgcatctgcgtccggcctctgagcttaaaatgtgacacagatgcaatgtgtatgtgaaacaagctc | 780  |
| Db | 3684 | tgtctgcatctgcgtccggcctctgagcttaaaatgtgacacagatgcaatgtgtatgtgaaacaagctc | 3743 |
| QY | 781  | tctgtccagtgctgtatgtgaatgcccacagaagaatcagttctatgtgcagtgtgattgtgaattggg  | 840  |
| Db | 3744 | tctgtccagtgctgtatgtgaatgcccacagaagaatcagttctatgtgcagtgtgattgtgaattggg  | 3803 |
| QY | 841  | caccacacctccggaagtgcacatcaatattctctgnaaaggtacaacatttctgtggaagaacca     | 900  |
| Db | 3804 | caccacacctccggaagtgcacatcaatattctctgnaaaggtacaacatttctgtggaagaacca     | 3863 |
| QY | 901  | tccgcagagcgctctcttggaatatctgcgcaataaacttcccttactgtctcgaacaacctgtgat    | 960  |
| Db | 3864 | tccgcagagcgctctcttggaatatctgcgcaataaacttcccttactgtctcgaacaacctgtgat    | 3923 |
| QY | 961  | ggacacttggaagcttctctactgttttctgtcatatctctctcccaacaaatgatagtgcatgga     | 1020 |
| Db | 3924 | ggacacttggaagcttctctactgttttctgtcatatctctctcccaacaaatgatagtgcatgga     | 3983 |

|    |      |  |      |
|----|------|--|------|
| QY | 1021 | agattatctccaagtagagcaagctgtccagaggaaccccaactacgaatgaaataataga    | 1080 |
| Db | 3984 | agcttaatgtccaagtagagcaagctgtccagaggaaccccaactacgaatgaaataataga   | 4043 |
| QY | 1081 | agaagcggagagactatgatagtatctactgattctgaaatgagatggtgcagattga       | 1140 |
| Db | 4044 | agaagcggagagactatgatagtatctactgattctgaaatgagatggtggtcagattga     | 4103 |
| QY | 1141 | tgatgaaactctctctctcttattccaattcgcctcaattgccaagaagatctctaaac      | 1200 |
| Db | 4104 | tgatgaaactctctctctcttattccaattcgcctcaattgccaagaagatctctaaac      | 4163 |
| QY | 1201 | ttgggtacatcacattgctgtcgtgaagaggaagactggtgactatgctcccttagtccgc    | 1260 |
| Db | 4164 | ttgggtacatcacattgctgtcgtgaagaggaagactggtgactatgctcccttagtccgc    | 4223 |
| QY | 1261 | ccccgacgacagaaagtataaaagtcataatttgacaacatggtccctcagcggattgtag    | 1320 |
| Db | 4224 | ccccgacgacagaaagtataaaagtcataatttgacaacatggtccctcagcggattgtag    | 4283 |
| QY | 1321 | gaagtagcaaaaaagtcggatttatgcatataacagataaaaacctttaagactgttgaagc   | 1380 |
| Db | 4284 | gaagtagcaaaaaagtcggatttatgcatataacagataaaaacctttaagactgttgaagc   | 4343 |
| QY | 1381 | tattcagaatgatatcagaatctcttgagaccttacttattatggtggaagtgtgagacacact | 1440 |
| Db | 4344 | tattcagaatgatatcagaatctcttgagaccttacttattatggtggaagtgtgagacacact | 4403 |
| QY | 1441 | gttgattatatttgaagatcaagcaagcagacacataacatacctacccctcaggaattac    | 1500 |
| Db | 4404 | gttgattatatttgaagatcaagcaagcagacacataacatacctacccctcaggaattac    | 4463 |
| QY | 1501 | tgaatccggtcccttgatataccaagagattcccaaaagtggttaaaccattggaagattt    | 1560 |
| Db | 4464 | tgaatccggtcccttgatataccaagagattcccaaaagtggttaaaccattggaagattt    | 4523 |
| QY | 1561 | tccaattctgcagaagagaataattccaatatanaatlygacaglygactgtagaagatggcc  | 1620 |
| Db | 4524 | tccaattctgcagaagagaataattccaatatanaatlygacaglygactgtagaagatggcc  | 4583 |
| QY | 1621 | aactaatccagaatccctgtgtgtccgtgcacccgcctattactcgtgttgcgttaatgagaa  | 1680 |
| Db | 4584 | aactaatccagaatccctgtgtgtccgtgcacccgcctattactcgtgttgcgttaatgagaa  | 4643 |
| QY | 1681 | agatcagcttccaagaccatctggtgcctctctcctacctcgtccacaagaagatctgtagtca | 1740 |
| Db | 4644 | agatcagcttccaagaccatctggtgcctctctcctacctcgtccacaagaagatctgtagtca | 4703 |
| QY | 1741 | aagaggaaccagataatgtcagacagaagaatgatactcctgttttctgtatttgatga      | 1800 |
| Db | 4704 | aagaggaaccagataatgtcagacagaagaatgatactcctgttttctgtatttgatga      | 4763 |
| QY | 1801 | gaaccggaagctgtgatccctcacagaagaataatacaagcttctccccaatccagctggaggt | 1860 |
| Db | 4764 | gaaccggaagctgtgatccctcacagaagaataatacaagcttctccccaatccagctggaggt | 4823 |
| QY | 1861 | gcagctttagagatccagaatttccaagcctccaacatactgacacagatcaatgtgtatgt   | 1920 |
| Db | 4824 | gcagctttagagatccagaatttccaagcctccaacatactgacacagatcaatgtgtatgt   | 4883 |
| QY | 1921 | tttgatagtttgcaagttgtgaatttgttttgatgaggttggtcactggtatattctaa      | 1980 |
| Db | 4884 | tttgatagtttgcaagttgtgaatttgttttgatgaggttggtcactggtatattctaa      | 4943 |
| QY | 1981 | catltagcacacagactgactccctctcgtctctctctcctcgtgatataccctcaaacaa    | 2040 |
| Db | 4944 | catltagcacacagactgactccctctcgtctctctctcctcgtgatataccctcaaacaa    | 5003 |
| QY | 2041 | aatggtctatagaagacacactcaacctatactccaattctcagggaaaactgtctcatgctc  | 2100 |
| Db | 5004 | aatggtctatagaagacacactcaacctatactccaattctcagggaaaactgtctcatgctc  | 5063 |

QY 2101 gatggaaaacccaggtctatgagttctgggtgctgcacaaactcgaagcttcgcgaacagag 2160  
|||||  
Db 5064 gatggaaaacccaggtctatgattctgggtgctgcacaaactcgaagcttcgcgaacagag 5123  
QY 2161 catgacgcgcctactgaaggttcttaagtgtgtaagaagaacctggttgatttaagagga 2220  
|||||  
Db 5124 catgacgcgcctactgaaggttcttaagtgtgtaagaagaacctggttgatttaagagga 5183  
QY 2221 cagttaagaagatacttcagacatactctgagttaaaaaaacatgcacattgaaccagaag 2280  
|||||  
Db 5184 cagttaagaagatacttcagacatactctgagttaaaaaaacatgcacattgaaccagaag 5243  
QY 2281 ctctccccaagaattcaagacacccctagacactagcgcaaaagaaattlaatgccccacc 2340  
|||||  
Db 5244 ctctccccaagaattcaagacacccctagacactagcgcaaaagaaattlaatgccccacc 5303  
QY 2341 agctcttaaaacgcatacaacgggaaataaactcgtactactcttcagtcagatcaagagga 2400  
|||||  
Db 5304 agctcttaaaacgcatacaacgggaaataaactcgtactactcttcagtcagatcaagagga 5363  
QY 2401 aattgaactatgatatccatccatcaagtgtgaatgaagaagaatttcgacattatga 2460  
|||||  
Db 5364 aattgaactatgatatccatccatcaagtgtgaatgaagaagaatttcgacattatga 5423  
QY 2461 tgaagatgaaaaatcaagagccccgcagccttcaaaagaaaaacgcacataattatg 2520  
|||||  
Db 5424 tgaagatgaaaaatcaagagccccgcagccttcaaaagaaaaacgcacataattatg 5483  
QY 2521 tgaagatgagaaggtctgggattatgggatagtagctcccaactgttcttaagaagaag 2580  
|||||  
Db 5484 tgaagatgagaaggtctgggattatgggatagtagctcccaactgttcttaagaagaag 5543  
QY 2581 ggtcacaagtgagcaggtgctccctcagttcaagaagaagtttctccagaattatcagtcg 2640  
|||||  
Db 5544 ggtcacaagtgagcaggtgctccctcagttcaagaagaagtttctccagaattatcagtcg 5603  
QY 2641 ctcccttaactcagccctatatacgttggaactaaatgaaacttgggaactcctgggccc 2700  
|||||  
Db 5604 ctcccttaactcagccctatatacgttggaactaaatgaaacttgggaactcctgggccc 5663  
QY 2701 atatatagaagcagaagtttgaagaataatacattgaacttcaagaatacggcctccg 2760  
|||||  
Db 5664 atatatagaagcagaagtttgaagaataatacattgaacttcaagaatacggcctccg 5723  
QY 2761 tccctatctcctcctatctcctatcttctatgagaagaatcgaagcgcaagcgaga 2820  
|||||  
Db 5724 tccctatctcctcctatctcctatcttctatgagaagaatcgaagcgcaagcgaga 5783  
QY 2821 acctagaaaaaacttgtcaagcctaatagaaacaaaacttacttgggaagtgcaca 2880  
|||||  
Db 5784 acctagaaaaaacttgtcaagcctaatagaaacaaaacttacttgggaagtgcaca 5843  
QY 2881 tcatatggcccccttaagaatgagtttgacttgaagaacggggtttttctcgaagt 2940  
|||||  
Db 5844 tcatatggcccccttaagaatgagtttgacttgaagaacggggtttttctcgaagt 5903  
QY 2941 tgaacctgaaaaaatagtgcactcagcgctgattggaaccccttcggtcgcacactaa 3000  
|||||  
Db 5904 tgaacctgaaaaaatagtgcactcagcgctgattggaaccccttcggtcgcacactaa 5963  
QY 3001 cacactgaacccctgtcactatggtgagaacagtgcagtaacagaatttgcctgttttcaac 6023  
|||||  
Db 3061 cacctttgattgagaccaaagctgttacttcaactgaaaaatttgaaaaaaactgagggc 6120  
|||||  
QY 6024 cacctttgattgagaccaaagctgttacttcaactgaaaaatttgaaaaaaactgagggc 6083  
|||||  
QY 3121 tcccttcgaatatccagatggaagatcccaactttaaagaagatatcccttccatgcagcat 3180  
|||||  
Db 6084 tcccttcgaatatccagatggaagatcccaactttaaagaagatatcccttccatgcagcat 6143  
QY 3181 caatgctacataatgatacactactgcttgaatggtcgaagtcaagaatgcg 3240  
|||||

Db 6144 caatgctacataatgatacactaacctggttgaatggtcgaagtcaagaatgcg 6203  
|||||  
QY 3241 atggtatctgtcgaagatgggcaagcaatgaaataatccattctatcttcaagtga 3300  
|||||  
Db 6204 atggtatctgtcgaagatgggcaagcaatgaaataatccattctatcttcaagtga 6263  
QY 3301 tgtgtcactgttacgaaaaaaggagtatataatgtgcaactgtacaaactctccagg 3360  
|||||  
Db 6264 tgtgtcactgttacgaaaaaaggagtatataatgtgcaactgtacaaactctccagg 6323  
QY 3361 tgttttgaagacagtggaaatggttaaccaaccgaactggaatttggcgggtggaatgct 3420  
|||||  
Db 6324 tgttttgaagacagtggaaatggttaaccaaccgaactggaatttggcgggtggaatgct 6383  
QY 3421 tatgtggagacatcaactgtctggatgagagacacttttctgtgtgacagaaatgtg 3480  
|||||  
Db 6384 tatgtggagacatcaactgtctggatgagagacacttttctgtgtgacagaaatgtg 6443  
QY 3481 tcaagatcccttgaggaaatggtctctggacacattagagatttccagattcaagcttcagg 3540  
|||||  
Db 6444 tcaagatcccttgaggaaatggtctctggacacattagagatttccagattcaagcttcagg 6503  
QY 3541 acaataatgacagtgggccccaaagctggccgaacttcaattatccggatccaatgac 3600  
|||||  
Db 6504 acaataatgacagtgggccccaaagctggccgaacttcaattatccggatccaatgac 6563  
QY 3601 ctggagcaccaagagaccccttctctgatacaaggttgatctgtgtgcacaaatgat 3660  
|||||  
Db 6564 ctggagcaccaagagaccccttctctgatacaaggttgatctgtgtgcacaaatgat 6623  
QY 3661 tcaacgcatcaagaaccagggtgcocgctcagaagtctccagcctcatalactccagtt 3720  
|||||  
Db 6624 tcaacgcatcaagaaccagggtgcocgctcagaagtctccagcctcatalactccagtt 6683  
QY 3721 tatcatcatgatagtctcttgatgggaagaagtggcagaacttatcgagaattccactgtg 3780  
|||||  
Db 6684 tatcatcatgatagtctcttgatgggaagaagtggcagaacttatcgagaattccactgtg 6743  
QY 3781 aaaccttaatgtctctcttggcaatgtgatactcctcgggataaaacacataattttaa 3840  
|||||  
Db 6744 aaaccttaatgtctctcttggcaatgtgatactcctcgggataaaacacataattttaa 6803  
QY 3841 cccctcaattatgtctcgatacatccggttggcacccaactcaattatagcatctgcagcac 3900  
|||||  
Db 6804 cccctcaattatgtctcgatacatccggttggcacccaactcaattatagcatctgcagcac 6863  
QY 3901 tcttcgcagtgaagtgtgaggtgtgatttaaatagttgcagcatctgggaatgga 3960  
|||||  
Db 6864 tcttcgcagtgaagtgtgaggtgtgatttaaatagttgcagcatctgggaatgga 6923  
QY 3961 gagtaagaacatatacagaatgcagagatgacttcaacttccacttaacaaatgtttgc 4020  
|||||  
Db 6924 gagtaagaacatatacagaatgcagagatgacttcaacttccacttaacaaatgtttgc 6983  
QY 4021 caactgtgtctccttcaaaagctgacttcaactcaccccaaggaggaatgacctgagagacc 4080  
|||||  
Db 6984 caactgtgtgtccttcaaaagctgacttcaactcaccccaaggaggaatgacctgagagacc 7043  
QY 4081 tcaagttgaataatccaaaagggtgctgcagaagtgtgacttccagaagaacatgaaagtccac 4140  
|||||  
Db 7044 tcaagttgaataatccaaaagggtgctgcagaagtgtgacttccagaagaacatgaaagtccac 7103  
QY 4141 aggaatcaactcagagagtaaaatcctcttaccagcagctgtatgtgaagggttccct 4200  
|||||  
Db 7104 aggaatcaactcagagagtaaaatcctcttaccagcagctgtatgtgaagggttccct 7163  
QY 4201 catctccagcatcaagaatggccaatcagttgagactctcttcttcaagaatggcaagttaa 4260  
|||||  
Db 7164 catctccagcatcaagaatggccaatcagttgagactctcttcttcaagaatggcaagttaa 7223  
QY 4261 ggttttcaagggaatcaagaactccttcaacactgtgtgaaactcttatacccacggt 4320  
|||||

Db 7224 gttttcagggaatacaagactcttcacacctgtgtgtaactctctagaccacggt 7283  
QY 4321 actgactcgtctacttgtaattacccccagagttgtgacacagattgcccgaagat 4380  
Db 7284 actgactcgtctacttgtaattacccccagagttgtgacacagattgcccgaagat 7343  
QY 4381 ggaagttctggcgtcgagacagagacctctactga 4417  
Db 7344 ggaagttctggcgtcgagacagagacctctactga 7380

RESULT 3  
AA076016  
ID AA076016 standard; cDNA; 4629 BP.  
XX  
AC AA076016;  
XX  
DT 20-JUL-1995 (first entry)  
XX  
DE B-domain deleted Factor-VIII.  
XX  
KW Factor-VIII; blood-clotting; hemophilia A; gene therapy;  
KW adenovirus; vector; ss.  
XX  
OS Homo sapiens.  
XX  
PN M09429471-A.  
XX  
XX 22-DEC-1994.  
XX  
XX 13-APR-1994; 94MO-US04075.  
XX  
PR 10-JUN-1993; 93US-0074920.  
PR 25-MAR-1994; 94US-0218355.  
XX  
XX (GENE-) GENETIC THERAPY INC.  
XX  
PI Connolly S, Kaleko M, Smith T;  
XX  
XX WPI; 1995-036495/05.  
DR P-PSDB; AAR67709.  
XX  
XX New adenoviral vectors for treatment of haemophilia - contg. a  
PT DNA sequence encoding a clotting factor, partic. Factor VIII or  
PT Factor IX  
XX  
PS Disclosure; Fig. 17A-17C; 116pp; English.  
XX  
XX Human Factor-VIII cDNA, from which the B domain had been deleted, was  
CC used to construct recombinant adenovirus vectors that produced  
CC therapeutic levels of the clotting factor when administered to an  
CC animal host, potentially providing hemophilia A gene therapy.  
XX  
SQ Sequence 4629 BP; 1318 A; 1051 C; 1018 G; 1242 T; 0 other;

Query Match 98.6%; Score 4355.2; DB 16; Length 4629;  
Best Local Similarity 99.1%; Pred. No. 0;  
Matches 4378; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 2 atcaaatagaagctccacactgtctcttctgtgctttgtcgattctgttagggc 61  
Db 1 atgcaaatagaagctccacactgtctcttctgtgctttgtcgattctgttagggc 60  
QY 62 accagaagatactactcgtgggtgcagtggaactgtcatgaggactatataatgcaagtgatctc 121  
Db 61 accagaagatactactcgtgggtgcagtggaactgtcatgaggactatataatgcaagtgatctc 120  
QY 122 ggtgagactgctgtgagcgaagaattccctccctagagtgccaaaattcttccatcaaac 181  
Db 121 ggtgagactgctgtgagcgaagaattccctccctagagtgccaaaattcttccatcaaac 180  
QY 182 acctcagctgtgataaaaagactctgtttagaattcaacggttcaacctttcaacatc 241

Db 181 acctcagctgtgataaaaagactctgtttagaattcaacggttcaacctttcaacatc 240  
QY 242 gctaagccaaggccaacctgtgatgtctgtgaagtcctacacacacagctgaagttat 301  
Db 241 gctaagccaaggccaacctgtgatgtctgtgaagtcctacacacacagctgaagttat 300  
QY 302 gatacaatgtgtatataactaagaacatgtctcccatctcctgtcaattctcaagctgtc 361  
Db 301 gatacaatgtgtatataactaagaacatgtctcccatctcctgtcaattctcaagctgtc 360  
QY 362 ggtgtatccctactggaagactcttcgaaggagctgaatatgatagtatcagaccagcaagg 421  
Db 361 ggtgtatccctactggaagactcttcgaaggagctgaatatgatagtatcagaccagcaagg 420  
QY 422 gagaagaagaatgataaagctctccctgtgtggaagccatacatatgtctgcgaagctctg 481  
Db 421 gagaagaagaatgataaagctctccctgtgtggaagccatacatatgtctgcgaagctctg 480  
QY 482 aaagaagatgtgccaatgacctgacctgaccactgtgaccttacctatcatcttctcat 541  
Db 481 aaagaagatgtgccaatgacctgacctgaccactgtgaccttacctatcatcttctcat 540  
QY 542 gtggaacctgtgtaaaagacttgaattcaggccctcatgtgagccctactagatgtagagaa 601  
Db 541 gtggaacctgtgtaaaagacttgaattcaggccctcatgtgagccctactactatgtagagaa 600  
QY 602 gggagctctggccaaggaaagacacagacctgtgacaaatttatactactttgtctgta 661  
Db 601 gggagctctggccaaggaaagacacagacctgtgacaaatttatactactttgtctgta 660  
QY 662 ttgtatgaagggaanaagtgtgcactcagaaacaaagaacctctgtgacggaataaggat 721  
Db 661 ttgtatgaagggaanaagtgtgcactcagaaacaaagaacctctgtgacggaataaggat 720  
QY 722 gctgcactctctcggtcggtcgccctaaatgcacacagctcaattgttataaagctct 781  
Db 721 gctgcactctctcggtcggtcgccctaaatgcacacagctcaattgttataaagctct 780  
QY 782 ctgccaagctcgtatgtgatgcacaggaatacagctcatatgtgcatgtgatatgggc 841  
Db 781 ctgccaagctcgtatgtgatgcacaggaatacagctcatatgtgcatgtgatatgggc 840  
QY 842 accactcctgaagtgcactcaatatctctcctgaaggtcaacacattcttgygaagaacct 901  
Db 841 accactcctgaagtgcactcaatatctctcctgaaggtcaacacattcttgygaagaacct 900  
QY 902 cgcagagctcccttggaatactcgcacaataacttcttactctgtcaaacacttgatg 961  
Db 901 cgcagagctcccttggaatactcgcacaataacttcttactctgtcaaacacttgatg 960  
QY 962 gaccttgaagattctactctgttctgcatatctctccacacacatgatatgcatgaa 1021  
Db 961 gaccttgaagattctactctgttctgcatatctctccacacacatgatatgcatgaa 1020  
QY 1022 gcttatgtcaaatgagacagctgtctcagaggaaccccaactagaatgaaataatgaa 1081  
Db 1021 gcttatgtcaaatgagacagctgtctcagaggaaccccaactagaatgaaataatgaa 1080  
QY 1082 ggaagcgaagaactatgatatgtatcctactacagattctgaatgtggtgtgcaagttgat 1141  
Db 1081 ggaagcgaagaactatgatatgtatcctactacagattctgaatgtggtgtgcaagttgat 1140  
QY 1142 gatgacaactctccctcttlatccaattcgtcagttgcgaagaagcactctaaact 1201  
Db 1141 gatgacaactctccctcttlatccaattcgtcagttgcgaagaagcactctaaact 1200  
QY 1202 tgggtaatatacatgtgtctgtaagaagagactggaactatgtctccctagctctggc 1261  
Db 1201 tgggtaatatacatgtgtctgtaagaagagactggaactatgtctccctagctctggc 1260  
QY 1262 ccgagtacagaagtataaagtcacatacttgacaacttgccctagcgagatgtgag 1321

Dh 1261 ccsgatgacagaagtltaanaagltcaatatttgaaacatgcccctcagcgagtggttagg 1320  
Qy 1322 aagfacaaaagtlccgatttataggtacacacagatgaaaccttaagaactgtaagct 1381  
Dh 1321 aagtaaaaaaagtlccgatttataggtacacacagatgaaaccttaagaactgtaagct 1380  
Qy 1382 atcagcatgaatcagaagactctggacacttaattatgggaagcttggagacacactg 1441  
Dh 1381 atcagcatgaatcagaagactctggacacttaattatgggaagcttggagacacactg 1440  
Qy 1442 ttgatttatattaagaatccaagcaagacacataataactaccctcaggaactcact 1501  
Dh 1441 ttgatttatattaagaatccaagcaagacacataataactaccctcaggaactcact 1500  
Qy 1502 gatgtccgtcccttgattatccaagagattacccaagaagtgtlaaaacatttgaagaattt 1561  
Dh 1501 gatgtccgtcccttgattatccaagagattacccaagaagtgtlaaaacatttgaagaattt 1560  
Qy 1562 ccaattctgccaagagaaatatccaataataaaggacagtgactgtagaagaatgggcca 1621  
Dh 1561 ccaattctgccaagagaaatatccaataataaaggacagtgactgtagaagaatgggcca 1620  
Qy 1622 actaatacagaatccctcgggtggtcgtgacccggtatattacttgatttgtaataatgagaga 1681  
Dh 1621 actaatacagaatccctcgggtggtcgtgacccggtatattacttgatttgtaataatgagaga 1680  
Qy 1682 gatctaagcttcaaggaactatctgacccctctcctcatctgtacaaaagaatctgtaagatcaa 1741  
Dh 1681 gatctaagcttcaaggaactatctgacccctctcctcatctgtacaaaagaatctgtaagatcaa 1740  
Qy 1742 agaggaacacagataatgtcagaacaagagaatgtcatcctgtttctgtatattgag 1801  
Dh 1741 agaggaacacagataatgtcagaacaagagaatgtcatcctgtttctgtatattgag 1800  
Qy 1802 aaccggaagctgtgtactcctacacagagaataataacaagctttctccccaatcccgctgagtg 1861  
Dh 1801 aaccggaagctgtgtactcctacacagagaataataacaagctttctccccaatcccgctgagtg 1860  
Qy 1862 cagcttgaggaatccaggaattccaaagcctccaacatcatgacagatcaatgtgcatgtt 1921  
Dh 1861 cagcttgaggaatccaggaattccaaagcctccaacatcatgacagatcaatgtgcatgtt 1920  
Qy 1922 ttgtatagttgtcagctgtgtcagttgtttgtgatagagtgagctactgtgtacattcctaaagc 1981  
Dh 1921 ttgtatagttgtcagctgtgtcagttgtttcctacacacgtgcactgtgtacattcctaaagc 1980  
Qy 1982 attgaggaacagagactgtactccttctgtcttctctctgtgataataactccaacacaaa 2041  
Dh 1981 attgaggaacagagactgtactccttctgtcttctctctgtgataataactccaacacaaa 2040  
Qy 2042 atggtctcatgaaagacacactacacctatcccatctcagagagaacactgtctcatgtcg 2101  
Dh 2041 atggtctcatgaaagacacactacacctatcccatctcagagagaacactgtctcatgtcg 2100  
Qy 2102 atggaacaaaccaggtctcatgtgattctgtgggtgtccacaactcagactcttcggaacagaagc 2161  
Dh 2101 atggaacaaaccaggtctcatgtgattctgtgggtgtccacaactcagactcttcggaacagaagc 2160  
Qy 2162 atgagcccttaacgaaaggttctcagttgtgacaagaacacagctgcatattatcagaagac 2221  
Dh 2161 atgagcccttaacgaaaggttctcagttgtgacaagaacacagctgcatattatcagaagac 2220  
Qy 2222 agtatagaagatatctcagcatactgtctgtgataaaaaacaatgcatctggaacccaagaagc 2281  
Dh 2221 agtatagaagatatctcagcatactgtctgtgataaaaaacaatgcatctggaacccaagaagc 2280  
Qy 2282 ttctccagaatccaagaacacctagcacatagggcaaaaagcaatttaattgtccacccaccca 2341  
Dh 2281 ttctccagaatccaagaacacctagcacatagggcaaaaagcaatttaattgtccacccaccca 2340  
Qy 2342 gtctctgaaacgcatcaacggaataaactcgtactactcttcagtctcagatcaagaagaa 2401  
Dh 2341 gtctctgaaacgcatcaacggaataaactcgtactactcttcagtctcagatcaagaagaa 2400

Qy 2402 atgacatgatatgataccataatcagttgaaatgaaagaagaatttgacatttatgat 2461  
Dh 2401 atgacatgatatgataccataatcagttgaaatgaaagaagaatttgacatttatgat 2460  
Qy 2462 gaggaatgaataatcagagcccccagacttccaagaagaacagacactatttatgtct 2521  
Dh 2461 gaggaatgaataatcagagcccccagacttccaagaagaacagacactatttatgtct 2520  
Qy 2522 gcaagtggagagctcttggaattatggaatggatctcccaactgttctaagaagaag 2581  
Dh 2521 gcaagtggagagctcttggaattatggaatggatctcccaactgttctaagaagaag 2580  
Qy 2582 gctcagagctgagaggtccctcagttccaagaagaagtgttttccaaggaattatcagctgagc 2641  
Dh 2581 gctcagagctgagaggtccctcagttccaagaagaagtgttttccaaggaattatcagctgagc 2640  
Qy 2642 tcccttactcagcccttaacccgtgagagaactaaatgacaatttggagactcctctgggcca 2701  
Dh 2641 tcccttactcagcccttaacccgtgagagaactaaatgacaatttggagactcctctgggcca 2700  
Qy 2702 tatataagagcagaagttgaaagataataatcatgtgtaacttccagaaatcagagctctcgt 2761  
Dh 2701 tatataagagcagaagttgaaagataataatcatgtgtaacttccagaaatcagagctctcgt 2760  
Qy 2762 cccatcttccctatctcagcccttaattcttataaggaagaatcagagagcaagagcaaa 2821  
Dh 2761 cccatcttccctatctcagcccttaattcttataaggaagaatcagagagcaagagcaaa 2820  
Qy 2822 cctagaaaaaacttgtcagaagctaaatgaacccaactactcttggaaagtgcacaat 2881  
Dh 2821 cctagaaaaaacttgtcagaagctaaatgaacccaactactcttggaaagtgcacaat 2880  
Qy 2882 catatggaccccaactaaagaatgtatgtactgctcagaagcctgtgttatctctcagatgt 2941  
Dh 2881 catatggaccccaactaaagaatgtatgtactgctcagaagcctgtgttatctctcagatgt 2940  
Qy 2942 gaccctggaacaaagatgtgtcactcagagcctgtatttgagccctctcgtgtccacactaac 3001  
Dh 2941 gaccctggaacaaagatgtgtcactcagagcctgtatttgagccctctcgtgtccacactaac 3000  
Qy 3002 acaactgaacccctgtcactatgagagacaagtgcacagttacagaagaattgtctgttttccacc 3061  
Dh 3001 acaactgaacccctgtcactatgagagacaagtgcacagttacagaagaattgtctgttttccacc 3060  
Qy 3062 atcttctgataggaccaaagcctgtgtacttcaactgaataataatggaagaacactggaagcct 3121  
Dh 3061 atcttctgataggaccaaagcctgtgtacttcaactgaataataatggaagaacactggaagcct 3120  
Qy 3122 cccctgcaatatccagaatggaagaatcccaactttaaagaagaattatcgcttccatgcaatc 3181  
Dh 3121 cccctgcaatatccagaatggaagaatcccaactttaaagaagaattatcgcttccatgcaatc 3180  
Qy 3182 aatggtctacataatggaatacactacccgtgtatgtatgtgtcagatccaagaattcga 3241  
Dh 3181 aatggtctacataatggaatacactacccgtgtatgtatgtgtcagatccaagaattcga 3240  
Qy 3242 tggatactgctcgaagcatgtggcagcaatgaaacatccaattcatctcatcttcagctggagcat 3301  
Dh 3241 tggatactgctcgaagcatgtggcagcaatgaaacatccaattcatctcatcttcagctggagcat 3300  
Qy 3302 gtgttcaactgttacaagaaaaagagagatataaaatggcactgttacaatctctatccaggt 3361  
Dh 3301 gtgttcaactgttacaagaaaaagagagatataaaatggcactgttacaatctctatccaggt 3360  
Qy 3362 gtttttgagacatgagaaatgttaccatccaagcttgaatttggcgggtggaatgtcctt 3421  
Dh 3361 gtttttgagacatgagaaatgttaccatccaagcttgaatttggcgggtggaatgtcctt 3420  
Qy 3422 attgaggagcatctacatgctcgggagatgagacacttttctcgggtatcagcaataagtg 3481  
Dh 3421 attgaggagcatctacatgctcgggagatgagacacttttctcgggtatcagcaataagtg 3480



|    |      |   |      |
|----|------|---|------|
| QY | 479  | ctgaagaagaaatggtccaaatggtccttgaaacccactgtgccccttaacccatactcttct     | 538  |
| Db | 504  | ctgaagaagaaatggtccaaatggtccttgaaacccactgtgccccttaacccatactcttct     | 563  |
| QY | 539  | catgtgagccttgtaaaagacttgaatcagaagcctcaatttgagccctacagatgttga        | 598  |
| Db | 564  | catgtgagccttgtaaaagacttgaatcagaagcctcaatttgagccctacagatgttga        | 623  |
| QY | 599  | gaagaggagctctgccaagaaagaaagacaaagaccccttgacaaattatactactttgtct      | 658  |
| Db | 624  | gaagaggagctctgccaagaaagaaagacaaagaccccttgacaaattatactactttgtct      | 683  |
| QY | 659  | gtattctgtatgaaggaaaaagtctggaactcagaacaaagaactcccttgatgtcagatag      | 718  |
| Db | 684  | gtattctgtatgaaggaaaaagtctggaactcagaacaaagaactcccttgatgtcagatag      | 743  |
| QY | 719  | gatgtctgatctgcgtccagagccttggtccctaaatgtgacacagttcaatgtttgtttaaaccag | 778  |
| Db | 744  | gatgtctgatctgcgtccagagccttggtccctaaatgtgacacagttcaatgtttgtttaaaccag | 803  |
| QY | 779  | tctccgccaaggtctctgtatgtgaatgccaacaggaatacagttctatgcaatgtgatatgt     | 838  |
| Db | 804  | tctccgccaaggtctctgtatgtgaatgccaacaggaatacagttctatgcaatgtgatatgt     | 863  |
| QY | 839  | ggcaccaactccttgaaagtggaactccaatattcctcgaaggtcaacacattctctgtgagaac   | 898  |
| Db | 864  | ggcaccaactccttgaaagtggaactccaatattcctcgaaggtcaacacattctctgtgagaac   | 923  |
| QY | 899  | catgcgcaagcgctcccttggaatactcgcgaataaacttcccttaactggttcaaaaccttg     | 958  |
| Db | 924  | catgcgcaagcgctcccttggaatactcgcgaataaacttcccttaactggttcaaaaccttg     | 983  |
| QY | 959  | atggaaccttggaacagtttctactagtcttctgtatatactcttcccaaccaatgtgaatg      | 1011 |
| Db | 984  | atggaaccttggaacagtttctactagtcttctgtatatactcttcccaaccaatgtgaatg      | 104  |
| QY | 1019 | gaagcttatgtccaaagtgaacagctgtgccagaagaaacccaactaagaatgtgaaaaaat      | 1077 |
| Db | 1044 | gaagcttatgtccaaagtgaacagctgtgccagaagaaacccaactaagaatgtgaaaaaat      | 110  |
| QY | 1079 | gaagaagcgggaagactatgagatgaaacttccgaacttccgaatgtgaatgtggtccagttt     | 1133 |
| Db | 1104 | gaagaagcgggaagactatgagatgaaacttccgaacttccgaatgtgaatgtggtccagttt     | 116  |
| QY | 1139 | gatgtatgaacaactcctctcccttatcccaaatctgcgtcagttgcgaagaagcaltccataa    | 1191 |
| Db | 1164 | gatgtatgaacaactcctctcccttatcccaaatctgcgtcagttgcgaagaagcaltccataa    | 122  |
| QY | 1199 | acttggtgaactatgaattgtcgtcgcgtgaagaaggaagacatgagcatatgtcccttaagcttc  | 125  |
| Db | 1224 | acttggtgaactatgaattgtcgtcgcgtgaagaaggaagacatgagcatatgtcccttaagcttc  | 128  |
| QY | 1259 | gcccccgatgaacagagttataaaagccaatatttgacaatgtgcccctcagcgagttgt        | 1311 |
| Db | 1284 | gcccccgatgaacagagttataaaagccaatatttgacaatgtgcccctcagcgagttgt        | 134  |
| QY | 1319 | aggaagtatacaaaaaagtcggatttataatgtgcatacagaatgaacaccttgaagactgtga    | 1377 |
| Db | 1344 | aggaagtatacaaaaaagtcggatttataatgtgcatacagaatgaacaccttgaagactgtga    | 140  |
| QY | 1379 | gctattcgaactgaatcagaagatcttggagaccttacttataatgtgggaagtgtgagacaa     | 1433 |
| Db | 1404 | gctattcgaactgaatcagaagatcttggagaccttacttataatgtgggaagtgtgagacaa     | 146  |
| QY | 1439 | ctgtgtgatatatttaaanaatacagaacagcagaccatataaccttcaacccctcaggaatc     | 1491 |
| Db | 1464 | ctgtgtgatatatttaaanaatacagaacagcagaccatataaccttcaacccctcaggaatc     | 152  |
| QY | 1499 | actgatgtccgtcccttctgtatctcaagaagattaccaaagggtgtaanaaatgtgaagat      | 1551 |
| Db | 1524 | actgatgtccgtcccttctgtatctcaagaagattaccaaagggtgtaanaaatgtgaagat      | 158  |

|    |      |  |      |
|----|------|--|------|
| QY | 1555 | ttccaatcttcgccaaggagaataattcaaatataaatgacagtgactgylaaagatgvg         | 1618 |
| Db | 1584 | ttccaatcttcgccaaggagaataattcaaatataaatgacagtgactgylaaagatgvg         | 1643 |
| QY | 1619 | ccaactaaataagatctccggtgactaacccgactataactctagtttcgttaatgtag          | 1678 |
| Db | 1644 | ccaactaaataagatctccggtgactaacccgactataactctagtttcgttaatgtag          | 1703 |
| QY | 1679 | agagatctagcttcaagagactaatctgacctctccatctcgtacaaagaatctgtagat         | 1738 |
| Db | 1704 | agagatctagcttcaagagactaatctgacctctccatctcgtacaaagaatctgtagat         | 1763 |
| QY | 1739 | caaaagaggaagaacccagataatgltcagaaagaagaatgtaactctgtttctgtattgat       | 1788 |
| Db | 1764 | caaaagaggaagaacccagataatgltcagaaagaagaagaatgtaactctgtttctgtattgat    | 1823 |
| QY | 1799 | gagaaacgaagaactgtgtacctacagaagaataatacaagcttctctcccaatccagcttga      | 1858 |
| Db | 1824 | gagaaacgaagaactgtgtacctacagaagaataatacaagcttctctcccaatccagcttga      | 1883 |
| QY | 1859 | gttcagagcttcgagagatcccaagattccaaagcttccaaatcatgacagagcatatgactat     | 1918 |
| Db | 1884 | gttcagagcttcgagagatcccaagattccaaagcttccaaatcatgacagagcatatgactat     | 1943 |
| QY | 1919 | gtttttgtatgtttcgagttgttgcaagttgtttgtcatgtaggtgacatacgtgatactcta      | 1978 |
| Db | 1944 | gtttttgtatgtttcgagttgttgcaagttgtttgtcatgtaggtgacatacgtgatactcta      | 2003 |
| QY | 1979 | agcatctgaaagacagagactgacatctctctgtctctctctcgtgataataccttcaaac        | 2038 |
| Db | 2004 | agcatctgaaagacagagactgacatctctctgtctctctctcgtgataataccttcaaac        | 2063 |
| QY | 2039 | aaaatggtgtctatgaaagacacactaaccttatcccatctctcagaggaactgttctctatg      | 2098 |
| Db | 2064 | aaaatggtgtctatgaaagacacactaaccttatcccatctctcagaggaactgttctctatg      | 2123 |
| QY | 2099 | tcgatvggaaaaacccaggtctatgtatctcgggtgtccacaactcagacttccgaacaga        | 2158 |
| Db | 2124 | tcgatvggaaaaacccaggtctatgtatctcgggtgtccacaactcagacttccgaacaga        | 2183 |
| QY | 2159 | ggcatgaccgcgtcttaccgtgaaggtttctagtgtgaaagaacaacgtgtaattatcagag       | 2218 |
| Db | 2184 | ggcatgaccgcgtcttaccgtgaaggtttctagtgtgaaagaacaacgtgtaattatcagag       | 2243 |
| QY | 2219 | gacagtttatgaaatatcttcagacatactgtgtctagtaaaagaatgccatgtaaacaga        | 2278 |
| Db | 2244 | gacagtttatgaaatatcttcagacatactgtgtctagtaaaagaatgccatgtaaacaga        | 2303 |
| QY | 2279 | agctctctccagaatltcaagaacacctagcactaggcaaaagaattaatgtccaccaca         | 2338 |
| Db | 2304 | agctctctccagaatltcaagaacacctagcactaggcaaaagaattaatgtccaccaca         | 2363 |
| QY | 2339 | ccagttcttggaaaacgcataacaggggaataatcctcgttactactcttcagtcagataagag     | 2398 |
| Db | 2364 | ccagttcttggaaaacgcataacaggggaataatcctcgttactactcttcagtcagataagag     | 2423 |
| QY | 2399 | gaaatgtgactatgtatcacataatcagttgtgaaatgaagaagaagaattgtgatactat        | 2458 |
| Db | 2424 | gaaatgtgactatgtatcacataatcagttgtgaaatgaagaagaagaattgtgatactat        | 2483 |
| QY | 2459 | gatgtagagatgtaaaaaatcagaagccccccgcagcttltcaaaagaagaacaacgacatacttatt | 2518 |
| Db | 2484 | gatgtagagatgtaaaaaatcagaagccccccgcagcttltcaaaagaagaacaacgacatacttatt | 2543 |
| QY | 2519 | gctgtcagttggaagagctctcgtgattatgggatagtagctctcccaacatgltctaaagaac     | 2578 |
| Db | 2544 | gctgtcagttggaagagctctcgtgattatgggatagtagctctcccaacatgltctaaagaac     | 2603 |
| QY | 2579 | agggctcgaagttgcatagtccctccagttcaaaaagaagttgttttccagaagaattactgat     | 2638 |
| Db | 2604 | agggctcgaagttgcatagtccctccagttcaaaaagaagttgttttccagaagaattactgat     | 2663 |
| QY | 2639 | ggctccatttaccgaccttatataccgtgtggagaactaaatgaaactttagagatccctggg      | 2698 |



Db 2664 ggcctcttactcagcccttataccgtgaggaactaaatgacaatttgggcctcttggg 2723  
QY 2699 ccaataataagaagcagaagtgtgaagataataatcagttaacttcagaatcaggcctc 2758  
Db 2724 ccaataataagaagcagaagtgtgaagataataatcagttaacttcagaatcaggcctc 2783  
QY 2759 gctccattactcctctatcttctagccttatttctataggaagttcagagcagaaggca 2818  
Db 2784 cgtccattactcctctatcttctagccttatttctataggaagttcagagcagaaggca 2843  
QY 2819 gaacctagaanaaaacttctgtaaacgccaatgaaacccaacttcttctgaaagtcaa 2878  
Db 2844 gaacctagaanaaaacttctgtaaacgccaatgaaacccaacttcttctgaaagtcaa 2933  
QY 2879 catcatatggcaccactaaagaatgagttgtaactgcaaaagcctgagcttattctctgat 2938  
Db 2904 catcatatggcaccactaaagaatgagttgtaactgcaaaagcctgagcttattctctgat 2963  
QY 2939 gttgacctggaanaaagtgtgcaactcagacctgatttgaaccttctggtctgccaact 2998  
Db 2964 gttgacctggaanaaagtgtgcaactcagacctgatttgaaccttctggtctgccaact 3023  
QY 2999 aacacactgaaacctgtctcatatggagaacaaagtgaacagtacaagaaattgtctgttttc 3058  
Db 3024 aacacactgaaacctgtctcatatggagaacaaagtgaacagtacaagaaattgtctgttttc 3083  
QY 3059 acaactcttatatgagaccaaagcctgttactcaactgaaatatgaaagaactgcag 3118  
Db 3084 acaactcttatatgagaccaaagcctgttactcaactgaaatatgaaagaactgcag 3143  
QY 3119 gctccctgcacatatccgaatggaagatcccaactttaaagaagatttgccttcacatga 3178  
Db 3144 gctccctgcacatatccgaatggaagatcccaactttaaagaagatttgccttcacatga 3203  
QY 3179 atcaatgtctacataatgatactacactcgtgcttaagttaatgctcagatccaagaagatc 3238  
Db 3204 atcaatgtctacataatgatactacactcgtgcttaagttaatgctcagatccaagaagatc 3263  
QY 3239 cgaatgtatctgtctcagcagctgagcagaatgaaacatcattctattctcagatgga 3298  
Db 3264 cgaatgtatctgtctcagcagctgagcagaatgaaacatcattctattctcagatgga 3323  
QY 3299 catgtgtcactgtacgaaaaaagaagagatataaaatggcactgtaacaactctatcca 3358  
Db 3324 catgtgtcactgtacgaaaaaagaagagatataaaatggcactgtaacaactctatcca 3383  
QY 3359 ggtgttttggagacagtgaatgttaccatccaagaacttggaatttggcggtggaatgc 3418  
Db 3384 ggtgttttggagacagtgaatgttaccatccaagaacttggaatttggcggtggaatgc 3443  
QY 3419 ctatattggcagaatctacatgtctggaatgagcaacttttctgtgtcagaagaataag 3478  
Db 3444 ctatattggcagaatctacatgtctggaatgagcaacttttctgtgtcagaagaataag 3503  
QY 3479 tgtcagaactccctgggaatggtccttctgacacatltagagatttccagatccagcttca 3538  
Db 3504 tgtcagaactccctgggaatggtccttctgacacatltagagatttccagatccagcttca 3563  
QY 3539 ggaacaataatgagacgtggccccaagaactgagcagacttcatatcccgatccaatcaat 3598  
Db 3564 ggaacaataatgagacgtggccccaagaactgagcagacttcatatcccgatccaatcaat 3623  
QY 3599 ggcctggagcaccagaagcccttcttctgatacaagttgagatctgttggcaccagaatc 3658  
Db 3624 ggcctggagcaccagaagcccttcttctgatacaagttgagatctgttggcaccagaatc 3683  
QY 3659 attcagcgcatcaagaccagcaggtgcccgtcagaagttccccaagccctcatatctcag 3718  
Db 3684 attcagcgcatcaagaccagcaggtgcccgtcagaagttccccaagccctcatatctcag 3743  
QY 3719 ttatcatcatgtatagttcttgaatgggaagaagtgcagacattacgaagaattccact 3778  
|||||

Db 3744 ttatcatcatgtatagttcttgaatgggaagaagtgcagacattatcgaggaaattccact 3803  
QY 3779 ggaaccttaatgtgtcttctcttcttggcaatgtgattcattcttgggataaaacaataatttc 3838  
Db 3804 ggaaccttaatgtgtcttctcttcttggcaatgtgattcattcttgggataaaacaataatttc 3863  
QY 3839 aaccttccaattatgtctgatacatatccggttttggacccaactctatatacatctcagc 3898  
Db 3864 aaccttccaattatgtctgatacatatccggttttggacccaactctatatacatctcagc 3923  
QY 3899 actcttcgcatggaatgtgagctggcgtgcatgtaaatatgctgcagacatgacatttggagatg 3958  
Db 3924 actcttcgcatggaatgtgagctggcgtgcatgtaaatatgctgcagacatgacatttggagatg 3983  
QY 3959 gagagtaagaacataatcagatgcaagatgaactgcttcaacttaccataatgatt 4018  
Db 3984 gagagtaagaacataatcagatgcaagatgaactgcttcaacttaccataatgatt 4043  
QY 4019 gccacctggtctccttcaaaaagctcgactccaactccaagaaggagtaatgctctgaga 4078  
Db 4044 gccacctggtctccttcaaaaagctcgactccaactccaagaaggagtaatgctctgaga 4103  
QY 4079 cctcaggtgataatcccaaaaagctgagctgcaagtgcacttccagaagacaatgaaagtc 4138  
Db 4104 cctcaggtgataatcccaaaaagctgagctgcaagtgcacttccagaagacaatgaaagtc 4163  
QY 4139 acagagtaactactcagggagtaaatctctgcttccagcagatgtatgtgaaggatc 4198  
Db 4164 acagagtaactactcagggagtaaatctctgcttccagcagatgtatgtgaaggatc 4223  
QY 4199 ctcaatccagcagatcaagatgagcactgagctcctcttcttcaagaatgcaagaatga 4258  
Db 4224 ctcaatccagcagatcaagatgagcactgagctcctcttcttcaagaatgcaagaatga 4283  
QY 4259 aaggttttcaagggaatccaagaactccttcaacactgtgttgaactctcagaaccacg 4318  
Db 4284 aaggttttcaagggaatccaagaactccttcaacactgtgttgaactctcagaaccacg 4343  
QY 4319 ttactgactcgctacacttctgaaatcaaccccaagatgtgggtgacccaagattgccttgaag 4378  
Db 4344 ttactgactcgctacacttctgaaatcaaccccaagatgtgggtgacccaagattgccttgaag 4403  
QY 4379 atgaggttctggcgtcgagcgacagaccttactga 4417  
Db 4404 atgaggttctggcgtcgagcgacagaccttactga 4442  
  
RESULT 5  
AAV19581  
ID AAV19581 standard; cDNA; 4832 BP.  
AC AAV19581;  
XX 06-AUG-1998 (first entry)  
XX  
DE Human factor VIII beta-domain deleted SQN deletion cDNA sequence.  
XX  
KW Replication defective; recombinant retrovirus; RRV; therapeutic protein;  
KW haemophilia; thrombosis; hypercoagulable disorder; liver disease; human;  
KW hepatitis; thalassemia; phenylketonuria; Lesch-Nyhan syndrome; diabetes;  
KW cystic fibrosis; Duchenne's Muscular Dystrophy; hypercholesterolemia;  
KW hypopituitarism; adenine deaminase deficiency; HIV infection; anemia;  
KW Guucher's syndrome; high blood pressure; Alzheimer's disease; autoimmune;  
KW inflammatory disease; factor VIII; ss.  
XX  
OS Homo sapiens.  
XX  
PN MO9800541-A2.  
XX  
PD 08-JUN-1998.  
XX  
PF 02-JUL-1997; 97MO-US11784.  
XX



PR 04-JUN-1997; 97US-0869309.  
PR 03-JUL-1996; 96US-0645601.  
PR 13-AUG-1996; 96US-0696381.  
XX  
XX (CHIR ) CHIRON CORP.  
XX Allen JR, Barber JR, Boder M, Chang SMW, Chong K;  
PI De LA VEGA D, DepolonJ, Greengard J, Hsu DC, Ibanez CE;  
PI Jolly DJ, Lee R, Mittelstaedt DM, Prussak CE, Respass JG;  
XX WPI: 1998-086966/08.  
DR P-PSDB; AAW46246.  
XX  
XX New replication defective recombinant retro-viruses - which can be  
PT administered to provide long term systemic expression of therapeutic  
PT protein in blood, useful in, e.g. treating hyper-coagulable  
PT disorders  
XX  
XX Example 28; Pages 210-213; 272pp; English.  
PS  
XX This CDNA encodes the beta-domain deleted SQN deletion protein of human  
CC factor VIII. This is used in the construction of recombinant retroviral  
CC vectors expressing human factor VIII. The invention provides the  
CC preparation of replication defective recombinant retrovirus (RVV)  
CC expressing a therapeutic protein. The RVV preparation is resistant to  
CC degradation by human complement and is capable of inducing long term  
CC systemic expression of the therapeutic protein when administered  
CC intravenously to a human. The long term systemic expression results in a  
CC measurable level of the therapeutic protein being produced in the blood  
CC of the human for a period of at least 30 days after the administration of  
CC the RVV vector preparation. RVV's can be used for in vivo delivery of  
CC therapeutic protein to treat, e.g. haemophilia A, haemophilia B,  
CC thrombosis, hypercoagulable disorders, liver diseases such as hepatitis,  
CC disorders such as thalassemia, phenylketonuria, Lesch-Nyhan syndrome,  
CC severe combined immunodeficiency (SCID), cystic fibrosis, Duchenne's  
CC muscular dystrophy, inherited emphysema, familial hypercholesterolemia,  
CC diabetes, hypopituitarism, adenine deaminase deficiency, alpha1-  
CC antitrypsin deficiency, Gaucher's syndrome, anaemia, infections such as  
CC HIV infection, high blood pressure, Alzheimer's disease, autoimmune or  
CC inflammatory disease or graft versus host disease. RVV's are capable of  
CC surviving inactivation in human serum thereby allowing efficient gene  
CC transfer over prolonged periods of time.  
XX  
XX Sequence 4832 BP; 1381 A; 1089 C; 1055 G; 1307 T; 0 other;  
SQ

Query Match 97.8%; Score 4321.4; DB 19; Length 4832;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 4374; Conservative 0; Mismatches 1; Indels 42; Gaps 1;

QY 1 catgcaaatagagctctccactgtctctctctctgtgcctcttgcatctgtcttagtcg 60  
DB 71 catgcaaatagagctctccactgtctctctctctgtgcctcttgcatctgtcttagtcg 130

QY 61 caccagaagatactactctgtgtgcagtggaactgtcatggaactatgcaaatgcatct 120  
DB 131 caccagaagatactactctgtgtgcagtggaactgtcatggaactatgcaaatgcatct 190

QY 121 cgggtgaactgctctgtgacgcaaatcttcctcctagaagtgccaanaattttccattaa 180  
DB 191 cgggtgaactgctctgtgacgcaaatcttcctcctagaagtgccaanaattttccattaa 250

QY 181 cactcagctcgtgtacaaaagactctgtcttgaatactcagtcacactttcaaat 240  
DB 251 cactcagctcgtgtacaaaagactctgtcttgaatactcagtcacactttcaaat 310

QY 241 cggtaagcacaagccacccttgatggtctgctaggtcctaccacccaagctgagttta 300  
DB 311 cggtaagcacaagccacccttgatggtctgctaggtcctaccacccaagctgagttta 370

QY 301 tgatacagtggtctattacacttaagaataagcttccatccctgcaacttcaagctgt 360  
DB 371 tgatacagtggtctattacacttaagaataagcttccatccctgcaacttcaagctgt 430

QY 361 tggatcctcacttgaaaagctctctgaaggagctgaatatgatcatcagaccagttcaag 420  
DB 431 tggatcctcacttgaaaagctctctgaaggagctgaatatgatcatcagaccagttcaag 490

QY 421 ggaagaagaatgataaagctctccctgtgtggaagccatacatatgtctgcaggtcct 480  
DB 491 ggaagaagaatgataaagctctccctgtgtggaagccatacatatgtctgcaggtcct 550

QY 481 gaaagaagaatgctcaatgagctctgaccacacttgcttaacctaccatctcttca 540  
DB 551 gaaagaagaatgctcaatgagctctgaccacacttgcttaacctaccatctcttca 610

QY 541 tgtggaacttgtaaaagacttgaaatccaggtcattatggagccctactagatgtaga 600  
DB 611 tgtggaacttgtaaaagacttgaaatccaggtcattatggagccctactagatgtaga 670

QY 601 aggaagctctgcccagaagaaaagacacagacctgacacaatttactacttttctgt 660  
DB 671 aggaagctctgcccagaagaaaagacacagacctgacacaatttactacttttctgt 730

QY 661 attgaagaaggaaaagttggcactcagaagaacagacctctgattgcaagataagga 720  
DB 731 attgaagaaggaaaagttggcactcagaagaacagacctctgattgcaagataagga 790

QY 721 tgtgcatctgctcgggcttggtcctaaatgacacagtcgaatggtatgtaaacaggtc 780  
DB 791 tgtgcatctgctcgggcttggtcctaaatgacacagtcgaatggtatgtaaacaggtc 850

QY 761 tctgcaggtctgattgtgacacaggaataatcagttctatgtgcagatgtgaatggg 840  
DB 831 tctgcaggtctgattgtgacacaggaataatcagttctatgtgcagatgtgaatggg 910

QY 841 caccctctcgaagtgacactcaatatctctcgaagtgacacactcttcttgaggaaaca 900  
DB 911 caccctctcgaagtgacactcaatatctctcgaagtgacacactcttcttgaggaaaca 970

QY 901 tcgcaagcgcgtcttggaatctcgcacaataacttcttacttgctcctaaacactctgat 960  
DB 971 tcgcaagcgcgtcttggaatctcgcacaataacttcttacttgctcctaaacactctgat 1030

QY 961 ggaacttggaagatttctactcgttttgtcatactcttccccaacaatgcatgcatgga 1020  
DB 1031 ggaacttggaagatttctactcgttttgtcatactcttccccaacaatgcatgcatgga 1090

QY 1021 agctatgtcaaatgatacagctgtccagaggaaccccaactaagaatgaaataatga 1080  
DB 1091 agctatgtcaaatgatacagctgtccagaggaaccccaactaagaatgaaataatga 1150

QY 1081 aagaagcgaagactatgatagtacttactgattctgaaatgatagtgtgcaggttga 1140  
DB 1151 aagaagcgaagactatgatagtacttactgattctgaaatgatagtgtgcaggttga 1210

QY 1141 tgatgacacactctctccttataccaatcgttcaggttgccaagaagatcctaaac 1200  
DB 1211 tgatgacacactctctccttataccaatcgttcaggttgccaagaagatcctaaac 1270

QY 1201 ttgggtacattacattgtgtctgtaagaaggaagcttggaactatctccctagtcgc 1260  
DB 1271 ttgggtacattacattgtgtctgtaagaaggaagcttggaactatctccctagtcgc 1330

QY 1261 ccccgatgacagaaggttataaaaatcaatattgaaacatgagccctcagagattgttag 1320  
DB 1331 ccccgatgacagaaggttataaaaatcaatattgaaacatgagccctcagagattgttag 1390

QY 1321 gaagtacaaaaaagtcgattatgtgcatacacagatgaaaccttaagactcgtgaagc 1380  
DB 1391 gaagtacaaaaaagtcgattatgtgcatacacagatgaaaccttaagactcgtgaagc 1450

QY 1381 tattcagatgaatcaggaactcttggaaccttacttattggggaaatttggaacacact 1440  
DB 1451 tattcagatgaatcaggaactcttggaaccttacttattggggaaatttggaacacact 1510

QY 1441 gtgtatataatttaagaatcaagcaagcaacacataaacatctacccctcaggaatcac 1500  
|||||  
Db 1511 gtgtatataatttaagaatcaagcaagcaacacataaacatctacccctcaggaatcac 1570  
QY 1501 tgaatgcgtcccttctgtatcttaagaagattaccaaaaggtgtataaacatttgaagattc 1560  
|||||  
Db 1571 tgaatgcgtcccttctgtatcttaagaagattaccaaaaggtgtataaacatttgaagattc 1630  
QY 1561 tccaatcttcgcagagagaataatccaataataatgagcaatgagcttgaagaatggggcc 1620  
|||||  
Db 1631 tccaatcttcgcagagagaataatccaataataatgagcaatgagcttgaagaatggggcc 1690  
QY 1621 aactaaatcagatccctcgtgtgcctgcagccgcctatactctggtctgttaataatgagag 1680  
|||||  
Db 1691 aactaaatcagatccctcgtgtgcctgcagccgcctatactctggtctgttaataatgagag 1750  
QY 1681 agatctagctcaggaactcaatctggccctccctcatctgctatacagaatctgtatgaca 1740  
|||||  
Db 1751 agatctagctcaggaactcaatctggccctccctcatctgctatacagaatctgtatgaca 1810  
QY 1741 aagagagaaacagataatgttcagagcaagagaatgtatccctgtttctgtatgttga 1800  
|||||  
Db 1811 aagagagaaacagataatgttcagagcaagagaatgtatccctgtttctgtatgttga 1870  
QY 1801 gaaccggaagctgtgaacctcaacagaaatatacaacgcttctccccaatccagcttgag 1860  
|||||  
Db 1871 gaaccggaagctgtgaacctcaacagaaatatacaacgcttctccccaatccagcttgag 1930  
QY 1861 gcaagctggaagatccagagatcccaagcctcccaacatcatgacagacatcaatgtgata 1920  
|||||  
Db 1931 gcaagctggaagatccagagatcccaagcctcccaacatcatgacagacatcaatgtgata 1990  
QY 1921 ttgtgataagtttgcaagttgcagttgtttgtgacagaggtggcactggtatcatcttaag 1980  
|||||  
Db 1991 ttgtgataagtttgcaagttgcagttgtttgtgacagaggtggcactggtatcatcttaag 2050  
QY 1981 catctgagcacaagactgaactcccttctgtctctctctcctgagatataccttcaaacaca 2040  
|||||  
Db 2051 catctgagcacaagactgaactcccttctgtctctctctcctgagatataccttcaaacaca 2110  
QY 2041 aatgtgctataagaacaacaactacacctaattcccaattccagagagaacatgtctcatgtc 2100  
|||||  
Db 2111 aatgtgctataagaacaacaactacacctaattcccaattccagagagaacatgtctcatgtc 2170  
QY 2101 gatggaagaaacccaggtctatgagatctcgtgggtgcacaaactcgaactcttcggaaacaga 2160  
|||||  
Db 2171 gatggaagaaacccaggtctatgagatctcgtgggtgcacaaactcgaactcttcggaaacaga 2230  
QY 2161 catgaccgccttaactgaagtttcttaagtctgtgaacagaacacatgtgtattatcagagga 2220  
|||||  
Db 2231 catgaccgccttaactgaagtttcttaagtctgtgaacagaacacatgtgtattatcagagga 2290  
QY 2221 caagtataagaataattcagacatactgtctgagttaaaaacaatgtcaatltgaaccaagaag 2280  
|||||  
Db 2291 caagtataagaataattcagacatactgtctgagttaaaaacaatgtcaatltgaaccaagaag 2350  
QY 2281 ctctccccaagaattcaagacacccctagcactagcagaacaaatcttaatgtccaccacac 2340  
|||||  
Db 2351 ctctccccaagaattcaagacacccctagcactagcagaacaaatcttaatgtccaccacac 2398  
QY 2341 agctctgaacacgcataacacgggaataaactcgtactactcttcagtcagatcaagaagga 2400  
|||||  
Db 2369 agctctgaacacgcataacacgggaataaactcgtactactcttcagtcagatcaagaagga 2428  
QY 2401 aatgtgactatgataccatcatcatcagttgaaatgaaagaagaagatttgaacatttga 2460  
|||||  
Db 2429 aatgtgactatgataccatcatcatcagttgaaatgaaagaagaagatttgaacatttga 2488  
QY 2461 tgaagatgaaataatcagaagcccccgcaagcttccaagaagaaaaacagacatactttatgtc 2520  
|||||  
Db 2489 tgaagatgaaataatcagaagcccccgcaagcttccaagaagaaaaacagacatactttatgtc 2548  
QY 2521 tgcagtgagaagagctcgtgggattatgagatgagtagtccccaatgttcttaagaacag 2580

Db 2549 tgcagtgagaagagctcgtgggattatgagatgagtagtccccaatgttcttaagaacag 2608  
|||||  
QY 2581 ggcctcaagatgagcagtgctccctcagctcaggaagaagtgcttctccaggaatttactgattg 2640  
|||||  
Db 2609 ggcctcaagatgagcagtgctccctcagctcaggaagaagtgcttctccaggaatttactgattg 2668  
QY 2641 ctcccttaaccagccctataacgttgagagactaaatgaacatttggactccctggggcc 2700  
|||||  
Db 2669 ctcccttaaccagccctataacgttgagagactaaatgaacatttggactccctggggcc 2728  
QY 2701 atataataagcagaagttgaagataaatacatgtgtaacttccagaataacagcctctcg 2760  
|||||  
Db 2729 atataataagcagaagttgaagataaatacatgtgtaacttccagaataacagcctctcg 2788  
QY 2761 tccctattcctcttactctagccttattcttcatatgagaagatcagaagcagaagcaga 2820  
|||||  
Db 2789 tccctattcctcttactctagccttattcttcatatgagaagatcagaagcagaagcaga 2848  
QY 2821 acctagaanaaactgttcaagccttaatgaaaccaaactacttgggaagtgcacaca 2880  
|||||  
Db 2849 acctagaanaaactgttcaagccttaatgaaaccaaactacttgggaagtgcacaca 2908  
QY 2881 tcatatgacccccaataaagatgagttgactgcagaagcctgggctatctctgtatgt 2940  
|||||  
Db 2909 tcatatgacccccaataaagatgagttgactgcagaagcctgggctatctctgtatgt 2968  
QY 2941 tgaactggaanaaagatgtgcacactaagcctgtatgtgaacccctctgtctgcacactaa 3000  
|||||  
Db 2969 tgaactggaanaaagatgtgcacactaagcctgtatgtgaacccctctgtctgcacactaa 3028  
QY 3001 caacactgaacccctgtctatgtggaagacagtgacagatcaagaatttctgtctttcac 3060  
|||||  
Db 3029 caacactgaacccctgtctatgtggaagacagtgacagatcaagaatttctgtctttcac 3088  
QY 3061 catcttctgataagacccaagaagctgtgtaactcactgaanaaataatgaaagaacatgcagggc 3120  
|||||  
Db 3089 catcttctgataagacccaagaagctgtgtaactcactgaanaaataatgaaagaacatgcagggc 3148  
QY 3121 tccctgcataatccagatgtgaagaatcccaactttaaagaagataatcgcctccatgcacat 3180  
|||||  
Db 3149 tccctgcataatccagatgtgaagaatcccaactttaaagaagataatcgcctccatgcacat 3208  
QY 3181 caatggtctacataatgataatgatacactacctgtgtgataatggtcagaatccaaagattcg 3240  
|||||  
Db 3209 caatggtctacataatgataatgatacactacctgtgtgataatggtcagaatccaaagattcg 3268  
QY 3241 atggtatctgtcagcattggtgcagcaatgaaacacatccattctatcttcagtgaca 3300  
|||||  
Db 3269 atggtatctgtcagcattggtgcagcaatgaaacacatccattctatcttcagtgaca 3328  
QY 3301 tgtgttcaactgtacgaaaaaagagagataaaatggtgcactgtacacatctctacag 3360  
|||||  
Db 3329 tgtgttcaactgtacgaaaaaagagagataaaatggtgcactgtacacatctctacag 3388  
QY 3361 tgtttttggaacagtggaataatgttaccataccaaagcgtgaatttgggggtgtgaatgcct 3420  
|||||  
Db 3389 tgtttttggaacagtggaataatgttaccataccaaagcgtgaatttgggggtgtgaatgcct 3448  
QY 3421 tatgtgcgaacatctacatctcgtgagatgacacacatttctcgtgtgacagcaataatgt 3480  
|||||  
Db 3449 tatgtgcgaacatctacatctcgtgagatgacacacatttctcgtgtgacagcaataatgt 3508  
QY 3481 tcaagactccctcgtggaatgtgctctcgtgacacatlaagatttccaagtccaagcttcag 3540  
|||||  
Db 3509 tcaagactccctcgtggaatgtgctctcgtgacacatlaagatttccaagtccaagcttcag 3568  
QY 3541 acataatgagcagtggtggccccaagcgtgcagagacttcattctccggatccaatcaatgc 3600  
|||||  
Db 3569 acataatgagcagtggtggccccaagcgtgcagagacttcattctccggatccaatcaatgc 3628  
QY 3601 ctggaagaccaagaagcccttctctgatacaaggtgagatctgttcgacacaaatgattac 3660  
|||||

```

Db 3629 ctggagcacaaggagcccttcttggatcaagtgatctgtggaaccaatgattat 3688
Qy 3661 tcaagcacaacaaggagtgcccgctcagaagcttccagccctacatctcaagt 3720
Db 3689 tcaagcacaacaaggagtgcccgctcagaagcttccagccctacatctcaagt 3748
Qy 3721 tatcatcatgatatgcttggatggagaagaatgagacatacgaagaaatccactgg 3780
Db 3749 tatcatcatgatatgcttggatggagaagaatgagacatacgaagaaatccactgg 3808
Qy 3781 aaccttaatgtcttcttggcaatgtagatcattcattggaataaacaatttttaa 3840
Db 3809 aaccttaatgtcttcttggcaatgtagatcattcattggaataaacaatttttaa 3868
Qy 3841 cccccaattatctgcgatacatcgcgttgcaccaactcattatagatcgcagcac 3900
Db 3869 cccccaattatctgcgatacatcgcgttgcaccaactcattatagatcgcagcac 3928
Qy 3901 tcttcgcaatggagtgatggcgtgctgatttaaatagttgcagcaatggagaaatgga 3960
Db 3929 tcttcgcaatggagtgatggcgtgctgatttaaatagttgcagcaatggagaaatgga 3988
Qy 3961 gagtaagaacatatacagatgcacagatattacgtctcattaccataatgtttgc 4020
Db 3989 gagtaagaacatatacagatgcacagatattacgtctcattaccataatgtttgc 4048
Qy 4021 caccgtgtctcctcacaagaagctgcacccccaaggaggaatgacctgagagac 4080
Db 4049 caccgtgtctcctcacaagaagctgcacccccaaggaggaatgacctgagagac 4108
Qy 4081 tcaagtgtaataatccaaagaagtgctgcaagtggaacttccagaagaatgaagtaac 4140
Db 4109 tcaagtgtaataatccaaagaagtgctgcaagtggaacttccagaagaatgaagtaac 4168
Qy 4141 agggagtaactcctaaggagtaaaatctctgtcttaacagcagatgtatgtgaaggaattcct 4200
Db 4169 agggagtaactcctaaggagtaaaatctctgtcttaacagcagatgtatgtgaaggaattcct 4228
Qy 4201 catctccagcagtcacaagtgcacatcagtggaactcttcttcagaaatggaagtaaa 4260
Db 4229 catctccagcagtcacaagtgcacatcagtggaactcttcttcagaaatggaagtaaa 4288
Qy 4261 ggttttccaaggaaatcaagaactccttcaacactgtgtgaactcttcagaccacggt 4320
Db 4289 ggttttccaaggaaatcaagaactccttcaacactgtgtgaactcttcagaccacggt 4348
Qy 4321 actgaactcgaactcttcaaatcccccagagtggtgacagagattgcacctgaagat 4380
Db 4349 actgaactcgaactcttcaaatcccccagagtggtgacagagattgcacctgaagat 4408
Qy 4381 ggaaggtctctggctgcagagcagaggaactctactga 4417
Db 4409 ggaaggtctctggctgcagagcagaggaactctactga 4445

```

```

RESULT 6
AAV15338
ID AAV15338 standard; DNA; 4832 BP.
AC AAV15338;
XX
DT 20-JUL-1998 (first entry)
XX
DE Human Factor VIII SQN deletion mutant DNA.
XX
KM Factor VIII; blood clotting; haemophilia A; gene therapy.
XX
KW retrovirus; vector; human; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
CDS 72..4445
FT

```

```

FT XX /*tag= a
PN XX MO3800542-A2.
PD XX 08-JAN-1998.
XX
PF 02-JUL-1997; 97WO-US11785.
XX
PR 04-JUN-1997; 97US-0869309.
PR 03-JUL-1996; 96US-0645601.
PR 13-AUG-1996; 96US-0696381.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Allen JR, Barber JR, Boder M, Chang SM, Chong K;
PI De la Vega D, Depolo NJ, Greengard J, Hsu DC, Ibanez CE;
PI Jolly DJ, Mittelstaedt DM, Prussak CE, Respass JG;
XX
DR WP1: 1998-086967/08.
DR P-PSDB; AAMW4372.
XX
PT New replication defective recombinant retroviruses - which express B
PT domain-deleted human factor VIII or human factor IX for the
PT treatment of haemophilia
XX
PS Claim 6; Page 174-175; 236pp; English.
XX
CC This DNA sequence includes a coding region for the B domain
CC deletion mutant SQN (see AAMW4372) of human Factor VIII. The SQN
CC mutant is created by fusing Ser-743 to Gln-1638 of native Factor
CC VIII (see AAMW4373) to form a Ser-Gln-Asn (SQN) link between the A2
CC and A3 Factor VIII domains. When compared to plasmid-derived
CC Factor VIII, the SQN deletion does not influence the in vivo
CC pharmacokinetics, but the reduced size of the molecule appears to
CC decrease proteolytic degradation. The invention relates to
CC preparations of replication defective recombinant retrovirus (RV)
CC expressing a B domain-deleted human Factor VIII protein, where the
CC recombinant RV is capable of infecting human cells, is resistant
CC to degradation by human complement and is capable of inducing
CC long-term (at least 30 days and up to 6 months or longer
CC post-injection) systemic expression of Factor VIII when
CC administered to a haemophilia A patient.
XX
SQ Sequence 4832 BP; 1381 A; 1089 C; 1055 G; 1307 T; 0 other;

```

```

Query Match 97.8%; Score 4321.4; DB 19; Length 4832;
Best Local Similarity 99.0%; Pred. No. 0; Mismatches 1; Indels 42; Gaps 1;
Matches 4374; Conservative 0;
Qy 1 catgcaaatagagctctccacctgcttcttctgtgtgcttggcattcgtcttagtc 60
Db 71 catgcaaatagagctctccacctgcttcttctgtgtgcttggcattcgtcttagtc 130
Qy 61 caccagaagatactactctggtgagtggaactgtcatggaactatataatgaagtatct 120
Db 131 caccagaagatactactctggtgagtggaactgtcatggaactatataatgaagtatct 190
Qy 121 cgggagctgctgtgagcgaagaattctcctcagaagtgccaaatctttcatctaa 180
Db 191 cgggagctgctgtgagcgaagaattctcctcagaagtgccaaatctttcatctaa 250
Qy 181 cactcaatctgtacaaaagaactctgttgtgaatcaacggttcaactttcaacat 240
Db 251 cactcaatctgtgtacaaaagaactctgttgtgaatcaacggttcaactttcaacat 310
Qy 241 cgttaagcagaagccaccctgagatggtctgctagatcctcaacatcagagctgagttta 300
Db 311 cgttaagcagaagccaccctgagatggtctgctagatcctcaacatcagagctgagttta 370
Qy 301 tgatacagtggtcattacacttaagaacaatgagcttccaccctgtgagctttagtcgt 360
Db 371 tgatacagtggtcattacacttaagaacaatgagcttccaccctgtgagctttagtcgt 430

```

QY 361 tgggtatctactggaagcttctgaggagctgaaatgatgatcatcagaccagtcgaaag 420  
|||||  
Db 431 tgggtatctactggaagcttctgaggagctgaaatgatgatcatcagaccagtcgaaag 430  
QY 421 gggagaaagatgatgaaagtctctccctgggtggaaagcatatcatgctcggagctcct 480  
|||||  
Db 491 gggagaaagatgatgaaagtctctccctgggtggaaagcatatcatgctcggagctcct 550  
QY 481 gaaagagaatggtcccaatgagctctgacccacgtgctccttaactacatctctctca 540  
|||||  
Db 551 gaaagagaatggtcccaatgagctctgacccacgtgctccttaactacatctctctca 610  
QY 541 tgggagacctggtaaagaagcttgaaatcagagctcattggagccctactagtatgtgaga 600  
|||||  
Db 611 tgggagacctggtaaagaagcttgaaatcagagctcattggagccctactagtatgtgaga 670  
QY 601 agggagctctggccaaggaagaaagacacagacctgtgcacaaatttatactacttttgcgt 660  
|||||  
Db 671 agggagctctggccaaggaagaaagacacagacctgtgcacaaatttatactacttttgcgt 730  
QY 661 atttgatgaagggaagaaagtgtgcacccagaacaaagaactccttgatgcagataggga 720  
|||||  
Db 731 atttgatgaagggaagaaagtgtgcacccagaacaaagaactccttgatgcagataggga 790  
QY 721 tgcgtcacctgcgtccgggacctggcctaaatgycacacagtcacatggttatgttaaagctgc 780  
|||||  
Db 791 tgcgtcacctgcgtccgggacctggcctaaatgycacacagtcacatggttatgttaaagctgc 850  
QY 781 tctgcagagctctgatctgagatgcacaggaatcatgcttatgtgcagatggaatggag 840  
|||||  
Db 851 tctgcagagctctgatctgagatgcacaggaatcatgcttatgtgcagatggaatggag 910  
QY 841 caccacacctggaagtgcactccaatatctctcgaaaggtcacacattctctgtgagagacca 900  
|||||  
Db 911 caccacacctggaagtgcactccaatatctctcgaaaggtcacacattctctgtgagagacca 970  
QY 901 tcccgagagctctcttggaatctcgcgcaataactttccctactgctgcacaaactcttgat 960  
|||||  
Db 971 tcccgagagctctcttggaatctcgcgcaataactttccctactgctgcacaaactcttgat 1030  
QY 961 ggaaccttgagacagtttctactggttctgtcatatctctcccaaccaatgatgagcatgga 1020  
|||||  
Db 1031 ggaaccttgagacagtttctactggttctgtcatatctctcccaaccaatgatgagcatgga 1090  
QY 1021 agcttatgctcaaaagttagacagctgtccagaggaaccccaactcagatgaaaaataatga 1080  
|||||  
Db 1091 agcttatgctcaaaagttagacagctgtccagaggaaccccaactcagatgaaaaataatga 1150  
QY 1081 agagagcgaagacatgatgatcatcttactgattcttgaaatgagatggtgtcaggtttga 1140  
|||||  
Db 1151 agagagcgaagacatgatgatcatcttactgattcttgaaatgagatggtgtcaggtttga 1210  
QY 1141 tgaatgaacactctcctctcttcaatctgcctcagttgccaagaagatcctctaaac 1200  
|||||  
Db 1211 tgaatgaacactctcctctcttcaatctgcctcagttgccaagaagatcctctctaaac 1270  
QY 1201 ttgggtacatataatgctgcgttgaaagagagagctgggactatgctcccttaagtctctgc 1260  
|||||  
Db 1271 ttgggtacatataatgctgcgttgaaagagagagctgggactatgctcccttaagtctctgc 1330  
QY 1261 ccccgatgacagaagttataaagaagcaatatgtgaacaatggcccacagaggttggtga 1320  
|||||  
Db 1331 ccccgatgacagaagttataaagaagcaatatgtgaacaatggcccacagaggttggtga 1390  
QY 1321 gaagtcacaaaaagtcogatttatggcatatcacagatgaacaccttlaagactcgtgaagc 1380  
|||||  
Db 1391 gaagtcacaaaaagtcogatttatggcatatcacagatgaacaccttlaagactcgtgaagc 1450  
QY 1381 tattcagcatgaatcaggaatctctgggaccttactttaggggaagcttggaacacact 1440  
|||||  
Db 1451 tattcagcatgaatcaggaatctctgggaccttactttaggggaagcttggaacacacact 1510

QY 1441 gttcatatataatcaagaatcaagcagaccataataacatctaccctcaggaatcac 1500  
|||||  
Db 1511 gttcatatataatcaagaatcaagcagaccataataacatctaccctcaggaatcac 1570  
QY 1501 tgaatgcgtctcttgtatcaaggagatataccaagaaggtglaaaaacatttgaaggattt 1560  
|||||  
Db 1571 tgaatgcgtctcttgtatcaaggagatataccaagaaggtglaaaaacatttgaaggattt 1630  
QY 1561 tccaattctgcagaggaataatctcaataataatgagacgtgacgtgtagaagatggcc 1620  
|||||  
Db 1631 tccaattctgcagaggaataatctcaataataatgagacgtgacgtgtagaagatggcc 1690  
QY 1621 aactaaatcagatccctcgggtgccttgagcccgctatctactcagttcttgtaatatggag 1680  
|||||  
Db 1691 aactaaatcagatccctcgggtgccttgagcccgctatctactcagttcttgtaatatggag 1750  
QY 1681 agatctagcttcaggaatcattggccctctccatcatctgctacaagaatctgtagatca 1740  
|||||  
Db 1751 agatctagcttcaggaatcattggccctctccatcatctgctacaagaatctgtagatca 1810  
QY 1741 aagaggaacacagataatgtcagaacagaggaatgtcatctcttcttgatattatga 1800  
|||||  
Db 1811 aagaggaacacagataatgtcagaacagaggaatgtcatctcttcttgatattatga 1870  
QY 1801 gaacccgaagctgtgtacctcacaagaaatatacaacgcttctcccaatccagctggagt 1860  
|||||  
Db 1871 gaacccgaagctgtgtacctcacaagaaatatacaacgcttctcccaatccagctggagt 1930  
QY 1861 gcaagcttgagatccagaggttccaagcctccaacatcatgacagatcatgtgtatgt 1920  
|||||  
Db 1931 gcaagcttgagatccagaggttccaagcctccaacatcatgacagatcatgtgtatgt 1990  
QY 1921 ttcttgatgttttcagatgttcagttgtgtgtgtgcagtgaggtgagcatgtgtatcttaag 1980  
|||||  
Db 1991 ttcttgatgttttcagatgttcagttgtgtgtgtgcagtgaggtgagcatgtgtatcttaag 2050  
QY 1981 catgtgagcagacagactcctctctgtctctctcctgcgtgcataccttcaaacacaa 2040  
|||||  
Db 2051 catgtgagcagacagactcctctctgtctctctcctgcgtgcataccttcaaacacaa 2110  
QY 2041 aatggtctatgaaagacacactccaactatcccatctcagaggaacgtgtctcatgctc 2100  
|||||  
Db 2111 aatggtctatgaaagacacactccaactatcccatctcagaggaacgtgtctcatgctc 2170  
QY 2101 gatggaagaacccaggtctatggaattctggggtggtcccaactcagacttccggaacagag 2160  
|||||  
Db 2171 gatggaagaacccaggtctatggaattctggggtggtcccaactcagacttccggaacagag 2230  
QY 2161 catgacccgcttactgaaggtttctagttgtgacagaacaactggtgatattacagaga 2220  
|||||  
Db 2231 catgacccgcttactgaaggtttctagttgtgacagaacaactggtgatattacagaga 2290  
QY 2221 cagttatgaagatatttcagcatactgtcgtgagtaaaaacaatgcatatgnaaccaagag 2280  
|||||  
Db 2291 cagttatgaagatatttcagcatactgtcgtgagtaaaaacaatgcatatgnaaccaagag 2350  
QY 2281 ctcttcccgaagatccaagacaccttagcactagggcaaaagcaatttaatgccaacccacc 2340  
|||||  
Db 2351 ctcttcccgaagatccaagacaccttagcactagggcaaaagcaatttaatgccaacccacc 2368  
QY 2341 agtcttgaaacgcatcaacgaggaataataactcgttactactctcagtcagatcaagaagga 2400  
|||||  
Db 2369 agtcttgaaacgcatcaacgaggaataataactcgttactactctcagtcagatcaagaagga 2428  
QY 2401 aatgtactatgatgatcatatcatcagttgaaatgaaaggaagatcttgacatttatga 2460  
|||||  
Db 2429 aatgtactatgatgatcatatcatcagttgaaatgaaaggaagatcttgacatttatga 2488  
QY 2461 tgaaggtgaaaatcagagcccgagcgttcaaaaagaaaacaagcactattttatgtc 2520  
|||||  
Db 2489 tgaaggtgaaaatcagagcccgagcgttcaaaaagaaaacaagcactattttatgtc 2548  
QY 2521 tgcaggtgagagagcctctggatattatggaatgagtagctcccaacatgcttctaagaagag 2580

|||||  
Db 2549 tgcagctggagagcctcggagctta tggatgtagagctccacatgcttctaagaacag 2608  
OY 2581 ggcctcaagctggagctgctccctcagttcaagaagctgtttccaggattactgacg 2640  
Db 2609 ggcctcaagctggagctgctccctcagttcaagaagctgtttccaggattactgacg 2668  
OY 2641 ctccttactcagcctctataccgtggagaaactaaatgaacatttggagctctggggcc 2700  
Db 2669 ctccttactcagcctctataccgtggagaaactaaatgaacatttggagctctggggcc 2728  
OY 2701 atatatagagcgaagtgtgaagataataatcatgttaacttccagaatccagcctccg 2760  
Db 2729 atataagagcgaagtgtgaagataataatcatgttaacttccagaatccagcctccg 2788  
OY 2761 tccctatctctctatctcagcctatttctctagagaaatccagagcgaaggacga 2820  
Db 2789 tccctatctctctatctcagcctatttctctagagaaatccagagcgaaggacga 2848  
OY 2821 acctagaaaaaacttgttcaagcctaa tgaaccaaaacttacttctggaagtgcaca 2880  
Db 2849 acctagaaaaaacttgttcaagcctaa tgaaccaaaacttacttctggaagtgcaca 2908  
OY 2881 tctatggcccccataaagatgagttgactgcagaagcctgggcttaattctctgagt 2940  
Db 2909 tctatggcccccataaagatgagttgactgcagaagcctgggcttaattctctgagt 2968  
OY 2941 tgcacctggaanaagatgtgcactcaggccttgatggacccctctggtctgcacactaa 3000  
Db 2969 tgcacctggaanaagatgtgcactcaggccttgatggacccctctggtctgcacactaa 3028  
OY 3001 caacactgaaccctgtctcatctggagacaaatgacagtaacagaatctgctgttttcac 3060  
Db 3029 caacactgaaccctgtctcatctggagacaaatgacagtaacagaatctgctgttttcac 3088  
OY 3061 catctcttgatgagaccaaaagctgttacttcaactgaaaaatctggaagaagaactgcagggc 3120  
Db 3089 catctcttgatgagaccaaaagctgttacttcaactgaaaaatctggaagaagaactgcagggc 3148  
OY 3121 tccctgcaatatccagatctggaagatcccaactttaaagagatattcctcctccatgcat 3160  
Db 3149 tccctgcaatatccagatctggaagatcccaactttaaagagatattcctcctccatgcat 3208  
OY 3181 caatggtctacataatgatacactactgctgtagtaatggcgaagatacaaaagattcg 3240  
Db 3209 caatggtctacataatgatacactactgctgtagtaatggcgaagatacaaaagattcg 3268  
OY 3241 atggtatctgtcagcctggtggcagcaatgaaaacatcatcttctatcttcaatgtgaca 3300  
Db 3269 atggtatctgtcagcctggtggcagcaatgaaaacatcatcttctatcttcaatgtgaca 3328  
OY 3301 tgggttcaactgtaagaaaaaagaggagataaaaatgacactgtcacatctctaccag 3360  
Db 3329 tgggttcaactgtaagaaaaaagaggagataaaaatgacactgtcacatctctaccag 3388  
OY 3361 tgggttctgagacagtggaatggttaacatccaagctggaattggcggtgtgaaatgctc 3420  
Db 3389 tgggttctgagacagtggaatggttaacatccaagctggaattggcggtgtgaaatgctc 3448  
OY 3421 tatgtggagcattctacatgctggtgagacacacttttctgtgtacacgaataagt 3460  
Db 3449 tatgtggagcattctacatgctggtgagacacacttttctgtgtacacgaataagt 3508  
OY 3481 tcaagactccctgggaatggcttctgacacatagagattcagatttaagcttcagg 3540  
Db 3509 tcaagactccctgggaatggcttctgacacatagagattcagatttaagcttcagg 3568  
OY 3541 acaatatgacagctgggcccacaaagctggcagactcatatctcgatcaatcaatgctc 3600  
Db 3569 acaatatgacagctgggcccacaaagctggcagactcatatctcgatcaatcaatgctc 3628  
OY 3601 ctggagaccaaagagcccttctctgatacaagtgatctgttggaccaaagtata 3660  
|||||

Db 3629 ctggagaccaaagagcccttctctgatacaagtgatctgttggcaccaatgattat 3688  
OY 3661 tcaaggacataagaccagggtgcccttcaagaatcttccagcctcacatctcagt 3720  
Db 3689 tcaaggacataagaccagggtgcccttcaagaatcttccagcctcacatctcagt 3748  
OY 3721 tatcatgtatagcttgaatgggaagaatggcagacttatccagaaattccactg 3780  
Db 3749 tatcatgtatagcttgaatgggaagaatggcagacttatccagaaattccactg 3808  
OY 3781 aaccttaatgcttctcttggcaatgfygattcaatctgggaataaacaataattttta 3840  
Db 3809 aaccttaatgcttctcttggcaatgfygattcaatctgggaataaacaataattttta 3868  
OY 3841 cccctcaattatctcgatataccgcttggacccaacatlataga ttcgcagcac 3900  
Db 3869 cccctcaattatctcgatataccgcttggacccaacatlataga ttcgcagcac 3928  
OY 3901 tcttcgcagtgaatgtagtggtgtgatttaaatagttgcagcatgccaatgggaatgga 3960  
Db 3929 tcttcgcagtgaatgtagtggtgtgatttaaatagttgcagcatgccaatgggaatgga 3988  
OY 3961 gagtaagcaatatcagatgacagatgactgcttcaactctacttcaacaatgatttcg 4020  
Db 3989 gagtaagcaatatcagatgacagatgactgcttcaactctacttcaacaatgatttcg 4048  
OY 4021 caactggtctcctccaaagctgcacttcaactcaccacaaaggaggaatgtaatgctggagac 4080  
Db 4049 caactggtctcctccaaagctgcacttcaactcaccacaaaggaggaatgtaatgctggagac 4108  
OY 4081 tcaagtgaaataatccaaaagatggtctgcagaatgagacttccagaagaacaatgaagtac 4140  
Db 4109 tcaagtgaaataatccaaaagatggtctgcagaatgagacttccagaagaacaatgaagtac 4168  
OY 4141 aggaatgaactgaactgaaggagtaaatctcgtcttaccagaatgattgtgaaggagttcct 4200  
Db 4169 aggaatgaactgaactgaaggagtaaatctcgtcttaccagaatgattgtgaaggagttcct 4228  
OY 4201 catctcagcagataagaatggtccatcagtgactctcttcttccagaatgcaaaataaa 4260  
Db 4229 catctcagcagataagaatggtccatcagtgactctcttcttccagaatgcaaaataaa 4288  
OY 4261 ggttttccaagggaatacaagaactcctcaacactgvtgvtgaaactcctagaccacggt 4320  
Db 4289 ggttttccaagggaatacaagaactcctcaacactgvtgvtgaaactcctagaccacggt 4348  
OY 4321 actgactcgtactccttgaattcaaccccccagagtttgggtgcaccagattgcccctggagat 4380  
Db 4349 actgactcgtactccttgaattcaaccccccagagtttgggtgcaccagattgcccctggagat 4408  
OY 4381 ggaagttctgggtcgagggcacaggaactctactga 4417  
Db 4409 ggaagttctgggtcgagggcacaggaactctactga 4445  
|||||

RESULT 7  
AAx82259  
ID AAx82259 standard; cDNA; 9164 BP.  
XX  
AC AAx82259;  
XX  
DT 18-AUG-1999 (first entry)  
XX  
DE Beta-domain deleted Factor VIII protein encoding gene (uncorrected).  
XX  
KW Factor VIII protein; gene modification; gene therapy; clinical disorder;  
KW splicing pattern; RNA processing; gene regulation; beta-domain; human;  
XX  
OS Homo sapiens.  
XX  
PN W09929848-A1.  
XX

|  |   |
|--|---|
| PD   | 17-JUN-1999.  |
| XX   |   |
| Pf   | 25-NOV-1998; 98WO-US25354.  |
| XX   |   |
| PR   | 16-JAN-1998; 98US-0071596.  |
| PR   | 05-DEC-1997; 97US-0067614.  |
| XX   |   |
| PA   | (IMMUN-) IMMUNE RESPONSE CORP.  |
| PI   | Bidlingmayer S, Gonzales JEN, Ill CR, Yang CO;                            |
| XX   |   |
| DR   | WPI: 1999-385602/32.  |
| XX   | P-FSDB: AAY21675.   |
| PT   | Genes and vectors exhibiting increased expression and novel splicing      |
| PT   | patterns, useful for expression of, e.g. beta-domain deleted factor       |
| PT   | VIII  |
| PS   |   |
| XX   | Example 2; Page 79-89; 123pp; English.                                    |
| XX   |   |
| CC   | The invention describes novel genes and vectors exhibiting increased      |
| CC   | expression and novel splicing patterns. It provides a gene encoding a     |
| CC   | Factor VIII protein, that comprises one or more consensus or near         |
| CC   | consensus splice sites which have been corrected to increase expression.  |
| CC   | The method, DNA sequences and expression vectors can be used to increase  |
| CC   | the expression of a gene, especially a Factor VIII gene. Genes containing |
| CC   | modified 5' and/or 3' untranslated regions have optimized expression      |
| CC   | levels and tissue-specific expression. The methods are used for           |
| CC   | identification and correction of consensus splice sites, addition of      |
| CC   | introns, optimization of 5' and 3' untranslated regions and increase in   |
| CC   | cytoplasmic RNA accumulation. Hence the DNAs are useful in gene therapy   |
| CC   | to treat a clinical disorder, to study RNA processing and/or gene         |
| CC   | regulation. The present sequence represents an uncorrected version of     |
| CC   | beta-domain deleted factor VIII protein encoding gene sequence (construct |
| CC   | pCV-2). This was used to develop a new coding sequence for beta domain    |
| CC   | deleted factor VIII protein by correcting the consensus splice sites.     |
| XX   |   |
| SQ   | Sequence 9164 BP; 2451 A; 2170 C; 2081 G; 2462 T; 0 other;                |
| <hr/>  |   |
| Query Match      97.8%; Score 4321.4; DB 20; Length 9164;                  |   |
| Best Local Similarity    99.0%; Pred. No. 0;                               |   |
| Matches 4374; Conservative    0; Mismatches    1; Indels    42; Gaps    1. |   |
| OY   | 1 catgcaaatagagctccacacctgtcttcttgctgccttttgatcgttcgttaagtc 60            |
| DB   | 1005 catggaatatagagctccctcacctgtcttcttgctgccttttgatcgttcgttaagtc 1064     |
| OY   | 61 caccagaataactactcctgtggtcagtyggaactgcatcagtgcataatgcaaagtatct 120      |
| DB   | 1065 caccagaataactactcctgtggtcagtyggaactgcatcagtgcataatgcaaagtatct 1124   |
| OY   | 121 cggtgagctgctgtggcgcaagatttcctcctctagatgccaaaatctttccattcaa 180        |
| DB   | 1125 cggtagctgctgtggcgcaagatttcctcctctagatgccaaaatctttccattcaa 1184       |
| OY   | 181 caccctagctgtgacaagaagacctgtttgaaatccaaggttcaaccttttaact 240           |
| DB   | 1185 caacctagctgtgacaagaagacctgtttgaaatccaaggttcaaccttttttaact 1244       |
| OY   | 241 cgttaagccaaggcaccctggatggtgctctgtagtgcctaccatccaggtcgaagtta 300       |
| DB   | 1245 cgttaagccaaggcaccctggatggtgctctgtagtgcctaccatccaggtcgaagtta 1304     |
| OY   | 301 tgatacatgtgtcaattaactttaagaacaatggtctcccatcctgtcagttctcatgctgt 360    |
| DB   | 1305 tgatacatgtgtcaattaactttaagaacaatggtctcccatcctgtcagttctcatgctgt 1364  |
| OY   | 361 tgggtatcccactctggaagctctcggagggagctgaatatgtagtgtcgaacagcaag 420       |
| DB   | 1365 tgggtatcccactctggaagctctcggagggagctgaatatgtagtgtcgaacagcaag 1424     |
| OY   | 421 ggaagaaagaatgatataagttccctcctgtgtggaagcataataatgtctggcaagttcct 480    |

|    |      |  |      |
|----|------|--|------|
| Db | 1425 | ggaagaagaatgataaagcttcctccctggtggaagccataacatgctcgtgcaggtccct  | 1484 |
| Oy | 481  | gaaagaagatggtcccaatgagcctctgacccacgtgccttaactaatactctccca      | 540  |
| Db | 1485 | gaaagaagaatggtcccaatgagcctctgacccacgtgccttaactaatactctccca     | 1544 |
| Oy | 541  | tgttgacctggtlaaagaacttgaaattlaagggcctcatctggaacctactagtatgtaga | 600  |
| Db | 1545 | tgttgacctggtlaaagaacttgaaattlaagggcctcatctggaacctactagtatgtaga | 1604 |
| Oy | 601  | aggaggtcttgcccaaggaaaagacagaccttgacaaattatactacttttgcgt        | 660  |
| Db | 1605 | aggaggtcttgcccaaggaaaagacagaccttgacaaattatactacttttgcgt        | 1664 |
| Oy | 661  | atttataaagggaagaatttgacctagaaaacaaagaactccttgatgcagatagga      | 720  |
| Db | 1665 | atttataaagggaagaatttgacctagaaaacaaagaactccttgatgcagatagga      | 1724 |
| Oy | 721  | tgtctgacatctgcgggccttgccctaaatagcacacagttcaatggtatgtaaacaggtc  | 780  |
| Db | 1725 | tgtctgacatctgcgggccttgccctaaatagcacacagttcaatggtatgtaaacaggtc  | 1784 |
| Oy | 781  | tcctgcaggtctgatatggaatgcacagaaatcaagtcctatctgacatgtaatggaatgg  | 840  |
| Db | 1785 | tcctgcaggtctgatatggaatgcacagaaatcaagtcctatctgacatgtaatggaatgg  | 1844 |
| Oy | 841  | caccactcttcaaatgtgacatcaaatctccgcgaagttcaacaattctctgttgaagacca | 900  |
| Db | 1845 | caccactcttcaaatgtgacatcaaatctccgcgaagttcaacaattctctgttgaagacca | 1904 |
| Oy | 901  | tcgcacagcgctcccttggaatctgcgcaataacttccctactgctcacaacactctgat   | 960  |
| Db | 1905 | tcgcacagcgctcccttggaatctgcgcaataacttccctactgctcacaacactctgat   | 1964 |
| Oy | 961  | ggaccttggacagttcttactggtttgtcatatctctccaccacaacatgaltgcatgga   | 1020 |
| Db | 1965 | ggaccttggacagttcttactggtttgtcatatctctccaccacaacatgaltgcatgga   | 2024 |
| Oy | 1021 | agcttatgtcaaaatagaacagctgttccagaggaacccacactagaatgtaaaaaataga  | 1080 |
| Db | 2025 | agcttatgtcaaaatagaacagctgttccagaggaacccacactagaatgtaaaaaataga  | 2084 |
| Oy | 1081 | agaagcggaaagacatgatgatgatacttactagaattcttgaaatgtaatggttga      | 1140 |
| Db | 2085 | agaagcggaaagacatgatgatgatacttactagaattcttgaaatgtaatggttga      | 2144 |
| Oy | 1141 | tgatgacaactctccctcccttaccatccaaattgcctcagttgccaagaagacatctaaac | 1200 |
| Db | 2145 | tgatgacaactctccctcccttaccatccaaattgcctcagttgccaagaagacatctaaac | 2204 |
| Oy | 1201 | tttggttaataacttgcgtcgtgaagaaggagacgtggactatgctcccttagtctctgc   | 1260 |
| Db | 2205 | tttggttaataacttgcgtcgtgaagaaggagacgtggactatgctcccttagtctctgc   | 2264 |
| Oy | 1261 | ccccgatgacgaagttataaagatcaatttggacaatggccctcagcggaattgtag      | 1320 |
| Db | 2265 | ccccgatgacgaagttataaagatcaatttggacaatggccctcagcggaattgtag      | 2324 |
| Oy | 1321 | gaagtacaaaaaagctccgatttatggtcatcacagaatgaaccttgaagctgtgaagc    | 1380 |
| Db | 2325 | gaagtacaaaaaagctccgatttatggtcatcacagaatgaaccttgaagctgtgaagc    | 2384 |
| Oy | 1381 | tattcagatgtaacacagaattcttggaacttaacttatttggggaagtgtggagaacact  | 1440 |
| Db | 2385 | tattcagatgtaacacagaattcttggaacttaacttatttggggaagtgtggagaacact  | 2444 |
| Oy | 1441 | gttgatatatttaagaatcaagaacgaacacataataacatcctccccaaggaataac     | 1500 |
| Db | 2445 | gttgatatatttaagaatcaagaacgaacacataataacatcctccccaaggaataac     | 2504 |
| Oy | 1501 | tgatgtccgtctcttgttatccaaggagatccaacaaagtgtlaaacaatttgaagattt   | 1560 |



Db 2505 tgaatgcgtcccttcttgatataccaagagataccaagaagctgtaaaacatctgaagattc 2564  
QY 1561 tccaattctgcagaagaaatattccaataataatggaagtgacttgagagaagagggcc 1620  
Db 2565 tccaattctgcagaagaaatattccaataataatggaagtgacttgagagaagagggcc 2624  
QY 1621 aactaaatcagatcctcgctgagcccgatctactagtttcgttaataatgagagag 1680  
Db 2625 aactaaatcagatcctcgctgagcccgatctactagtttcgttaataatgagagag 2684  
QY 1681 agatctagtttcagactcaatctgcccctctctcatctgctacaagaatctgtatagca 1740  
Db 2685 agatctagtttcagactcaatctgcccctctctcatctgctacaagaatctgtatagca 2744  
QY 1741 aagaagaaacagaataatgtcagaagaaggaatgtacacccgttttctgtattgtatga 1800  
Db 2745 aagaagaaacagaataatgtcagaagaaggaatgtacacccgttttctgtattgtatga 2804  
QY 1801 gaaccgaagctggtacactcacagaagaatatacaacgcttctccccaatccagctgagat 1860  
Db 2805 gaaccgaagctggtacactcacagaagaatatacaacgcttctccccaatccagctgagat 2864  
QY 1861 gcaagcttgagatcccaagatcccaagctcccaatcatctgacagcaatctgtatagt 1920  
Db 2865 gcaagcttgagatcccaagatcccaagctcccaatcatctgacagcaatctgtatagt 2924  
QY 1921 ttctgatagtttgagctgtctagtttgtctgacagagtgagatactgtatcatctag 1980  
Db 2925 ttctgatagtttgagctgtctagtttgtctgacagagtgagatactgtatcatctag 2984  
QY 1981 catctgagacaagaactgaactcctctctgctctctctctctgagataatcctcoaacaca 2040  
Db 2985 catctgagacaagaactgaactcctctctgctctctctctctgagataatcctcoaacaca 3044  
QY 2041 aatggtctatgaagaacacactcacccctattcccatctccagagaagaatctgtctatctc 2100  
Db 3045 aatggtctatgaagaacacactcacccctattcccatctccagagaagaatctgtctatctc 3104  
QY 2101 gatcgaaacacacagcttatatgactctgagggtgcacacacacagactctcggaacagag 2160  
Db 3105 gatcgaaacacacagcttatatgactctgagggtgcacacacacagactctcggaacagag 3164  
QY 2161 catgacgcgcttactgagaagttcttagttgtgacaagaacacctggtatattaccagaga 2220  
Db 3165 catgacgcgcttactgagaagttcttagttgtgacaagaacacctggtatattaccagaga 3224  
QY 2221 caattagaagaataattcagaatactgtgtgataaagaacatgcatctggaacagag 2280  
Db 3225 caattagaagaataattcagaatactgtgtgataaagaacatgcatctggaacagag 3284  
QY 2281 ctctctcccaaatcaagaacacccctagacactagccaagaatcttaatgccccacac 2340  
Db 3285 ctctctcccaaga-----ccacac 3302  
QY 2341 agctctgaaacgcacatcaacgggaataaactcgtaactcttcaagtcagaatcaagaga 2400  
Db 3303 agctctgaaacgcacatcaacgggaataaactcgtaactcttcaagtcagaatcaagaga 3362  
QY 2401 aattgactatgataatccatccatatcagttgaaatgaagaagaagaatttgcacattaga 2460  
Db 3363 aattgactatgataatccatccatatcagttgaaatgaagaagaagaatttgcacattaga 3422  
QY 2461 tggagatgaataatcaagcccccgacgcttccaaagaacaacacacttttatctgc 2520  
Db 3423 tggagatgaataatcaagcccccgacgcttccaaagaacaacacacttttatctgc 3482  
QY 2521 tgcagtgagaagctctcggaattatggaatgagtagctccccaactgttctaagaacag 2580  
Db 3483 tgcagtgagaagctctcggaattatggaatgagtagctccccaactgttctaagaacag 3542  
QY 2581 ggtctcaagatggcaggtctccctcagtcagaagaaggtgttctcaggaattactcagag 2640  
Db 3543 ggtctcaagatggcaggtctccctcagtcagaagaaggtgttctcaggaattactcagag 3602

QY 2641 ctcccttactcagcccttataccgtggaagaactaatgaacatttggagctcctgaggcc 2700  
Db 3603 ctcccttactcagcccttataccgtggaagaactaatgaacatttggagctcctgaggcc 3662  
QY 2701 atataagaagcagaagttgaagaataataatggttaacttccaagaatcagccctcg 2760  
Db 3663 atataagaagcagaagttgaagaataataatggttaacttccaagaatcagccctcg 3722  
QY 2761 tcccatcctcttacttactagcattattcttatgaggaagatcaaggccaaggagaga 2820  
Db 3723 tcccatcctcttacttactagcattattcttatgaggaagatcaaggccaaggagaga 3782  
QY 2821 aactagaanaaacttctgcaagcctaataatgaacaaacttacttctgaaatgacaca 2880  
Db 3783 aactagaanaaacttctgcaagcctaataatgaacaaacttacttctgaaatgacaca 3842  
QY 2881 tcaatgaccccaataagaatgagttgactgcaagcctggctatctctatgt 2940  
Db 3843 tcaatgaccccaataagaatgagttgactgcaagcctggctatctctatgt 3902  
QY 2941 tgaacttgaaaaaagatgacactcaggtctgattggacccctctggtctgccaactaa 3000  
Db 3903 tgaacttgaaaaaagatgacactcaggtctgattggacccctctggtctgccaactaa 3060  
QY 3001 cacactgaaccctgctcaatgggaacaagtgaacagtaacagaatctgtcttctcac 3060  
Db 3963 cacactgaaccctgctcaatgggaacaagtgaacagtaacagaatctgtcttctcac 4022  
QY 3061 catctctgaatgagccaagaagctggtactctactgaaataatggaagaacacacagagc 3120  
Db 4023 catctctgaatgagccaagaagctggtactctactgaaataatggaagaacacacagagc 4082  
QY 3121 tccctgcaataccaagatggaagatcccaactttaaagaagaattatcgcttccatgcaat 3180  
Db 4083 tccctgcaataccaagatggaagatcccaactttaaagaagaattatcgcttccatgcaat 4142  
QY 3181 caatgctacataatgataacactacactcgtgcttagtaatggtctcaggaatccaagattcg 3240  
Db 4143 caatgctacataatgataacactacactcgtgcttagtaatggtctcaggaatccaagattcg 4202  
QY 3241 atggtatctgctcagcaatgggcagcaatgaaacaatccatctctatcttcaagtgagaca 3300  
Db 4203 atggtatctgctcagcaatgggcagcaatgaaacaatccatctctatcttcaagtgagaca 4262  
QY 3301 tgtgtcactgtacgaaaaaagaaggagatataaatggtcacgtgtaacatctctacag 3360  
Db 4263 tgtgtcactgtacgaaaaaagaaggagatataaatggtcacgtgtaacatctctacag 4322  
QY 3361 tgttttgaagacagtgaagaatgtttacacatccaagcttgaaatttgagcgtggaatgct 3420  
Db 4323 tgttttgaagacagtgaagaatgtttacacatccaagcttgaaatttgagcgtggaatgct 4382  
QY 3421 tatgvgcagacatctacatgctctggatgagacacactttctctggttgaacgaataatg 3480  
Db 4383 tatgvgcagacatctacatgctctggatgagacacactttctctggttgaacgaataatg 4442  
QY 3481 tcaagctccctctggaatggtctctgacacattaggaatttcaagattcaagttcagag 3540  
Db 4443 tcaagctccctctggaatggtctctgacacattaggaatttcaagattcaagttcagag 4502  
QY 3541 acaatatgagacatgvggcccccaagaagcttggccaagcttcaatctatctcggaatcaaatg 3600  
Db 4503 acaatatgagacatgvggcccccaagaagcttggccaagcttcaatctatctcggaatcaaatg 4562  
QY 3601 ctggaagcaaccaagagcccttctctggaatcaaggtgagatctgttggcacaacatgatat 3660  
Db 4563 ctggaagcaaccaagagcccttctctggaatcaaggtgagatctgttggcacaacatgatat 4622  
QY 3661 tcaaggaatcaagaacccaggttgcccgtaagaagttctccagcctctacatctctcaggt 3720  
Db 4623 tcaaggaatcaagaacccaggttgcccgtaagaagttctccagcctctacatctctcaggt 4682



QY 3721 tatcatcatgtatagctctgtatggaagaatggtgcagacttatccagaaatccactg 3780  
|||||  
Db 4683 tatcatcatgtatagctctgtatggaagaatggtgcagacttatccagaaatccactg 4742  
QY 3781 aaccttaatggtctctcttggcaatgtgtatcattcgtggataaacaacataattttaa 3840  
|||||  
Db 4743 aaccttaatggtctctcttggcaatgtgtatcattcgtggataaacaacataattttaa 4802  
QY 3841 cccctcaattatgtctgtatataccgtttgtcccccacattcattatgattcgcagcac 3900  
|||||  
Db 4803 cccctcaattatgtctgtatataccgtttgtcccccacattcattatgattcgcagcac 4862  
QY 3901 tcttcgcatgagtgatgagtggtctgtatttaataatgattgcagcatggtggaatgga 3960  
|||||  
Db 4863 tcttcgcatgagtgatgagtggtctgtatttaataatgattgcagcatggtggaatgga 4922  
QY 3961 gagtaagaacaatcagatgcagagatctactgcttccacttaccataatgattgc 4020  
|||||  
Db 4923 gagtaagaacaatcagatgcagagatctactgcttccacttaccataatgattgc 4982  
QY 4021 cacttggtcctcttcaaaaagctgacttccctcccaaggagagatgattcctcggagacc 4080  
|||||  
Db 4983 cacttggtcctcttcaaaaagctgacttccctcccaaggagagatgattcctcggagacc 5042  
QY 4081 tcagtgtaataatccaaaagaatggtctgcagtgacttccagaagaacatgaaatgcac 4140  
|||||  
Db 5043 tcagtgtaataatccaaaagaatggtctgcagtgacttccagaagaacatgaaatgcac 5102  
QY 4141 aggaagtaactactcagggagagtaaaatctctgtaccagatgattggaaggattcct 4200  
|||||  
Db 5103 aggaagtaactactcagggagagtaaaatctctgtaccagatgattggaaggattcct 5162  
QY 4201 catctccagagatcagatgagtcagatgacttcttcttccagaaatggaagaatgaa 4260  
|||||  
Db 5163 catctccagagatcagatgagtcagatgacttcttcttccagaaatggaagaatgaa 5222  
QY 4261 ggttttcaaggagaatcaaaactccttcacacactgtgtgaactctctagaccacggt 4320  
|||||  
Db 5223 ggttttcaaggagaatcaaaactccttcacacactgtgtgaactctctagaccacggt 5282  
QY 4321 actgaactgcctactcctcgaattcaaccccaagattgggtgcacacagattgcccctgagat 4380  
|||||  
Db 5283 actgaactgcctactcctcgaattcaaccccaagattgggtgcacacagattgcccctgagat 5342  
QY 4381 ggaagttctggtctggaagcagagacactctactga 4417  
|||||  
Db 5343 ggaagttctggtctggaagcagagacactctactga 5379

RESULT 8  
AAD00122  
ID AAD00122 standard; DNA: 4999 BP.  
XX  
AC AAD00122;  
XX  
DT 31-JUL-2000 (first entry)  
XX  
DE Recombinant adeno associated vector construct, pVmA.1cF8deltaB.  
XX  
XX Recombinant Adeno Associated Vector; rAAV; 1cF8deltaB; promoter;  
KW human Factor VIII; hFVIII; EFlalpha; human elongation factor-1alpha;  
KW human growth hormone; hGH; ITR; inverted terminal repeat; haemophilia;  
XX gene therapy; ds.  
XX  
OS Adeno associated virus.  
XX  
PN WO200023116-A1.  
XX  
PD 27-APR-2000.  
XX  
PF 19-OCT-1999; 99WO-US24495.  
XX  
PR 20-OCT-1998; 98US-0104994.

PR 24-MAR-1999; 99US-0125974.  
PR 30-JUL-1999; 99US-0364862.  
XX  
PA (AVIG-) AVIGEN INC.  
XX  
PI Couto LB, Colosi PC;  
XX  
DR WPI; 2000-339536/29.  
XX  
XX New recombinant adenovirus-associated vector, useful for gene therapy  
PT to treat hemophilia, comprises at least a portion of Factor VIII  
PT operably linked to control sequence -  
XX  
PS Example 9; Fig 6; 92pp: English.  
XX  
CC The present DNA sequence is a recombinant adenovirus-associated vector,  
CC (rAAV) construct pVmA.1cF8deltaB. This expression vector comprises the  
CC promoter, first intron (-573 to +985) of human elongation factor-1alpha  
CC (EFlalpha) gene, human Factor VIII coding sequence (hFVIII) and a  
CC polyadenylation signal from human growth hormone (hGH). This sequence is  
CC inserted between the AAV inverted terminal repeat (ITR) regions. The  
CC hFVIII coding region comprises the heavy chain gene segment with the  
CC first 57 bp encoding the 19 amino acid signal peptide, A1 and A2 domains  
CC and 5 amino acids from the N-terminus of the B domain. The light chain  
CC segment comprises the C-terminal 85 amino acids of B domain and the A3,  
CC C1 and C2 domains. Both the heavy and light chain segments are cloned  
CC into the same plasmid separated by 42 nucleotides coding for 14 residues  
CC of the B domain, that is deleted. This plasmid is operably linked to  
CC control sequences, that directs the transcription and translation of the  
CC Factor VIII gene. The adeno-associated viral vectors are used for gene  
CC therapy to treat haemophilia. This method allows prolonged expression of  
CC therapeutic levels of Factor VIII in vivo. The rAAV are used for gene  
CC therapy, because of their broad host range, safety profile and duration  
CC of expression in the infected hosts.  
XX  
SQ Sequence 4999 BP; 1375 A; 1151 C; 1155 G; 1318 T; 0 other;

Query Match 97.8%; Score 4319.8; DB 21; Length 4999;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 4373; Conservative 0; Mismatches 2; Indels 42; Gaps 1;

QY 1 catcaaatagagctctccacactgctctcttctgtgctcttggatctgctttagtc 60  
|||||  
Db 415 catgcaaatagagctctccacactgctctcttctgtgctcttggatctgctttagtc 474  
QY 61 caccagaagatactacctggtggtcagtggaactgtcatggaactatgcaaaagtatct 120  
|||||  
Db 475 caccagaagatactacctggtggtcagtggaactgtcatggaactatgcaaaagtatct 534  
QY 121 cgtgtgaactgctctgtggaagcaaatcttctcctcagaagtgccaaatctttccattcaa 180  
|||||  
Db 535 cgtgtgaactgctctgtggaagcaaatcttctcctcagaagtgccaaatctttccattcaa 594  
QY 181 cactcagtcgtgtacaaaagaactctgtgtgaatcaacggttcaactttcaacat 240  
|||||  
Db 595 cactcagtcgtgtacaaaagaactctgtgtgaatcaacggttcaactttcaacat 654  
QY 241 cgtctaagcacaagggccaccctgagatggtctgtcctagtcctaccatccaggctgagttta 300  
|||||  
Db 655 cgtctaagcacaagggccaccctgagatggtctgtcctagtcctaccatccaggctgagttta 714  
QY 301 tgatacagtggtcattacacttaagaacatggtctcccatccctgctcagttctatgctgt 360  
|||||  
Db 715 tgatacagtggtcattacacttaagaacatggtctcccatccctgctcagttctatgctgt 774  
QY 361 tgggtatcctactcaggaagctcttgaggagctggaatgatgatcatcagaccaaag 420  
|||||  
Db 775 tgggtatcctactcaggaagctcttgaggagctggaatgatgatcatcagaccaaag 834  
QY 421 ggaagaagaatgataaagcttccctggttggaagcacaatatagtctgacaggtcct 480  
|||||  
Db 835 ggaagaagaatgataaagcttccctggttggaagcacaatatagtctgacaggtcct 894

|    |      |  |      |
|----|------|--|------|
| QY | 481  | gaagagaaatggttccaaatggtcctctgacccacgtgtgccttaacctatactcttccca      | 540  |
| Db | 895  | gaaagaaatggttccaaatggtcctctgacccacgtgtgccttaacctatactcttccca       | 954  |
| QY | 541  | tgtgagccttgtaaagaacttgtaattgaagcctatgttgagcccttaccagatgttagga      | 600  |
| Db | 955  | tgtgagccttgtaaagaacttgtaattgaagcctatgttgagcccttaccagatgttagga      | 1014 |
| QY | 601  | aggagatcttgcccaagaaaagacacagaccttgacaaatttatactatttgcgt            | 660  |
| Db | 1015 | aggagatcttgcccaagaaaagacacagaccttgacaaatttatactatttgcgt            | 1074 |
| QY | 661  | atttgatgaaggaaaagtggtgcactgaagaacaaagaacctcttgatlgcagataagga       | 720  |
| Db | 1075 | atttgatgaaggaaaagtggtgcactgaagaacaaagaacctcttgatlgcagataagga       | 1134 |
| QY | 721  | tgttgatctgtgtccggccctgtgacctaaatgacacagtcgaatgtgttgaacaggtc        | 780  |
| Db | 1135 | tgttgatctgtgtccggccctgtgacctaaatgacacagtcgaatgtgttgaacaggtc        | 1194 |
| QY | 781  | tctgcagagtcgtatltggaatgcacaaagaaatccagtcattgtgcagatggaatggtg       | 840  |
| Db | 1195 | tctgcagagtcgtatltggaatgcacaaagaaatccagtcattgtgcagatggaatggtg       | 1254 |
| QY | 841  | caacgaactccgaaggtgcaccccaataatctcccgaaaggtccacacattctctgtgaggaacca | 900  |
| Db | 1255 | caacgaactccgaaggtgcaccccaataatctcccgaaaggtccacacattctctgtgaggaacca | 1314 |
| QY | 901  | tccgaagcgctcccttggaatatctgcgcaataaacttcccttactgtcgtcaaacacttgat    | 960  |
| Db | 1315 | tccgaagcgctcccttggaatatctgcgcaataaacttcccttactgtcgtcaaacacttgat    | 1374 |
| QY | 961  | ggaaccttggaagtttctactgttttgcatactcttcccaacaatgaatggtcatgga         | 1020 |
| Db | 1375 | ggaaccttggaagtttctactgttttgcatactcttcccaacaatggtcatgga             | 1434 |
| QY | 1021 | agcttatgttcaaaagtagacagcgtgtccagaagaaaccccaactacgaatgtaaaaaataga   | 1080 |
| Db | 1435 | agcttatgttcaaaagtagacagcgtgtccagaagaaaccccaactacgaatgtaaaaaataga   | 1494 |
| QY | 1081 | agaagaaggaaagactatgatatgatactctactgattccgaaatgtagtgggtcagaatttga   | 1140 |
| Db | 1495 | agaagaaggaaagactatgatatgatactctactgattccgaaatgtagtgggtcagaatttga   | 1554 |
| QY | 1141 | tgaatgaacactctctctcttataccaattgcgtcagttgccaagaagcatctctaaac        | 1200 |
| Db | 1555 | tgaatgaacactctctctcttataccaattgcgtcagttgccaagaagcatctctaaac        | 1614 |
| QY | 1201 | tgtggtacattataatgtgtgtcgaaggaggagactgtgagactgtcccttaagtctcgc       | 1260 |
| Db | 1615 | tgtggtacattataatgtgtgtcgaaggaggagactgtgagactgtcccttaagtctcgc       | 1674 |
| QY | 1261 | ccccgaatgaacagaattatataaagtcacataattgaacaatgtgcctcagcgaattgtgag    | 1320 |
| Db | 1675 | ccccgaatgaacagaattatataaagtcacataattgaacaatgtgcctcagcgaattgtgag    | 1734 |
| QY | 1321 | gaagtaacaaaaagtcgcattatgtgcatacacagatgaaacctttaagactgttgaagc       | 1380 |
| Db | 1735 | gaagtaacaaaaagtcgcattatgtgcatacacagatgaaacctttaagactgttgaagc       | 1794 |
| QY | 1381 | tatttcagcatgaatcagaagaaatcttggaaccttttaattatgtgggaagtgtgaacacact   | 1440 |
| Db | 1795 | tatttcagcatgaatcagaagaaatcttggaaccttttaattatgtgggaagtgtgaacacact   | 1854 |
| QY | 1441 | gttgaattatataatgaagaatccagaacagaacacataaaccttccacctccaggaattac     | 1500 |
| Db | 1855 | gttgaattatataatgaagaatccagaacagaacacataaaccttccacctccaggaattac     | 1914 |
| QY | 1501 | tgaatgtcgtccttgttatccaaggagattaccaaagggtgtaaacatttgaagattt         | 1560 |
| Db | 1915 | tgaatgtcgtccttgttatccaaggagattaccaaagggtgtaaacatttgaagattt         | 1974 |

|    |      |   |      |
|----|------|---|------|
| QY | 1561 | lccaatlctgcagagaaataatatccaataataatgacagagctcgttagaagaatggccc   | 1620 |
| Db | 1975 | lccaatlctgcagagagaaatatccaataataatgacagagctcgttagaagaatggccc    | 2034 |
| QY | 1621 | aactaaatcaatctccggtgcgcgcgcgcctatgactagttctcgtaatatgagag        | 1680 |
| Db | 2035 | aactaaatcaatctccggtgcgcgcgcgcctatgactagttctcgtaatatgagag        | 2094 |
| QY | 1681 | agatctagcttcagagactcaatctgcccctccctcattcgtctcaagaatcgttagatca   | 1740 |
| Db | 2095 | agatctagcttcagagactcaatctgcccctccctcattcgtctcaagaatcgttagatca   | 2154 |
| QY | 1741 | aagaggaacacagataatctgcagaacagaggaatctcatccctggtttctcgtatttgatga | 1800 |
| Db | 2155 | aagaggaacacagataatctgcagaacagaggaatctcatccctggtttctcgtatttgatga | 2214 |
| QY | 1801 | gaacccgaagctcgtgtaactccaacagaaataaacaagctcttccccaatccagctcgaggt | 1860 |
| Db | 2215 | gaacccgaagctcgtgtaactccaacagaaataaacaagctcttccccaatccagctcgaggt | 2274 |
| QY | 1861 | gcagcttgcagagatccaagagttccaagcctccaacatcatgacaagcatcaatgagtaagt | 1920 |
| Db | 2275 | gcagcttgcagagatccaagagttccaagcctccaacatcatgacaagcatcaatgagtaagt | 2334 |
| QY | 1921 | tttgcataagtttgcagttctgcagtttggttgcataaggtgcgcatactgtaatacttcaag | 1980 |
| Db | 2335 | tttgcataagtttgcagttctgcagtttggttgcataaggtgcgcatactgtaatacttcaag | 2394 |
| QY | 1981 | catggagacaaagcctgaactccctcttcgtctctctctctgtgatatactcttaacaacaa  | 2040 |
| Db | 2395 | catggagacaaagcctgaactccctcttcgtctctctctctgtgatatactcttaacaacaa  | 2454 |
| QY | 2041 | aatgctcatatgaagacacactccaacccattcccatctcagaagagaaactgcttcatgtc  | 2100 |
| Db | 2455 | aatgctcatatgaagacacactccaacccattcccatctcagaagagaaactgcttcatgtc  | 2514 |
| QY | 2101 | gatvgaaaaacccaggtctatgagatctctgggtgcacaactcagaactcttcgnaacagag  | 2160 |
| Db | 2515 | gatvgaaaaacccaggtctatgagatctctgggtgcacaactcagaactcttcgnaacagag  | 2574 |
| QY | 2161 | catgaacgccttactagaagtttctcagttgtgtgcagaagaaactcgttgatattaaagaga | 2220 |
| Db | 2575 | catgaacgccttactagaagtttctcagttgtgtgcagaagaaactcgttgatattaaagaga | 2634 |
| QY | 2221 | cagttatgaagaatacttccaacatactgcgcgtgaanaaacaatgcatatgaaaccaagag  | 2280 |
| Db | 2635 | cagttatgaagaatacttccaacatactgcgcgtgaanaaacaatgcatatgaaaccaagag  | 2694 |
| QY | 2281 | ctctccccaagaattcaagacacccctagcacctagcgaagaacaaatthaatgccaacccac | 2340 |
| Db | 2695 | ctctccccaagaattcaagacacccctagcacctagcgaagaacaaatthaatgccaacccac | 2712 |
| QY | 2341 | agctctgaagaagcacaacaagggaaataaactgtactactcttcagctcagatcaagaaga  | 2400 |
| Db | 2713 | agctctgaagaagcacaacaagggaaataaactgtactactcttcagctcagatcaagaaga  | 2772 |
| QY | 2401 | aattgacatgatgatalccatccatccatgtgaagaatgaaagagaaatttgcacattatga  | 2460 |
| Db | 2773 | aattgacatgatgatalccatccatccatgtgaagaatgaaagagaaatttgcacattatga  | 2832 |
| QY | 2461 | tgaagatgaanaatcaagagcccccgcagcttccaagaagaacaagacacattattatgc    | 2520 |
| Db | 2833 | tgaagatgaanaatcaagagcccccgcagcttccaagaagaacaagacacattattatgc    | 2892 |
| QY | 2521 | tgcacgtgagagagctctcgtggaattatgagttagcttcccacaatgttttaagaagaag   | 2580 |
| Db | 2893 | tgcacgtgagagagctctcgtggaattatgagttagcttcccacaatgttttaagaagaag   | 2952 |
| QY | 2581 | ggcccaagtgagaggtctccctcagttccaagaagaagtgtttccaggaatttaacagatg   | 2640 |
| Db | 2953 | ggcccaagtgagaggtctccctcagttccaagaagaagtgtttccaggaatttaacagatg   | 3012 |
| QY | 2641 | ctctcttactaagcctctataccgtgagaaacthaatgaaacttgcggagctcctcggccc   | 2700 |

|||||  
Db 3013 ctcccttaccagcccttataccgttgagaacataagacatttggactccgagcc 3072  
Qy 2701 atatatagaagcgaagtgtgaagataatcatcagtgactttagaataagcccttcg 2760  
Db 3073 atatatagaagcgaagtgtgaagataatcatcagtgactttagaataagcccttcg 3132  
Qy 2761 tccctattctctctatcttgacttattcttaagagaatcagaagcagaagcaga 2820  
Db 3133 tccctattctctctatcttgacttattcttaagagaatcagaagcagaagcaga 3192  
Qy 2821 acctagaaaaaacttctgcaagcctaaatgaacccaaacttacttctggaagtcgaca 2880  
Db 3193 acctagaaaaaacttctgcaagcctaaatgaacccaaacttacttctggaagtcgaca 3252  
Qy 2881 tcatatggcccccaataaagatgatttgaactgcaaaagccttggtctattctctgagt 2940  
Db 3253 tcatatggcccccaataaagatgatttgaactgcaaaagccttggtctattctctgagt 3312  
Qy 2941 tgaacctgaaaaagatgtgacaccagccgtgattggaaccccttcgtctgcacaataa 3000  
Db 3313 tgaacctgaaaaagatgtgacaccagccgtgattggaaccccttcgtctgcacaataa 3372  
Qy 3001 cacactgaaccctgcctcatgtgagacaaagtacagtaacaggaatttgcctgtttcac 3060  
Db 3373 cacactgaaccctgcctcatgtgagacaaagtacagtaacaggaatttgcctgtttcac 3432  
Qy 3061 catcttctgataagcacaagaagctgtaacttcaactgaaataatggaagaagcctgcaggc 3120  
Db 3433 catcttctgataagcacaagaagctgtaacttcaactgaaataatggaagaagcctgcaggc 3492  
Qy 3121 tccctccaatatccagatgagaatcccaattttaagaagaattatgccttccatgtaaat 3180  
Db 3493 tccctccaatatccagatgagaatcccaattttaagaagaattatgccttccatgtaaat 3552  
Qy 3181 caatggtctacataatggaatacacttgcctgtgtaagtaatgctcagaatcaaaagattcg 3240  
Db 3553 caatggtctacataatggaatacacttgcctgtgtaagtaatgctcagaatcaaaagattcg 3612  
Qy 3241 atggtctctgctagaatgagcagcaatgaaacatccatctattctattcttcagtgcaca 3300  
Db 3613 atggtctctgctagaatgagcagcaatgaaacatccatctattctattcttcagtgcaca 3672  
Qy 3301 tgtgttcaactgtacgaaaaaagaagagataaataatgagcacttacatctcatccagg 3360  
Db 3673 tgtgttcaactgtacgaaaaaagaagagataaataatgagcactgtacatctcatccagg 3732  
Qy 3361 tgtttttgagacagtggaatgtttacatccaaagctggaatttggcggtgggaatgcct 3420  
Db 3733 tgtttttgagacagtggaatgtttacatccaaagctggaatttggcggtgggaatgcct 3792  
Qy 3421 tatgtggagacatctacatgctggtgagacacacttttctcgtgtgtaacagaataatg 3480  
Db 3793 tatgtggagacatctacatgctggtgagacacacttttctcgtgtgtaacagaataatg 3852  
Qy 3481 tccagactccctctggagatggtctctgagacatctagaagatttccagattacagcttcagg 3540  
Db 3853 tccagactccctctggagatggtctctgagacatctagaagatttccagattacagcttcagg 3912  
Qy 3541 acaatatgagcagtggtgccccaaagctgtgccaagctcatctatctcggatcaatcaatgc 3600  
Db 3913 acaatatgagcagtggtgccccaaagctgtgccaagctcatctatctcggatcaatcaatgc 3972  
Qy 3601 ctgtgagcacaagaagccttctctgtgatacgaagtgagatctgtgtgacccaatgatctac 3660  
Db 3973 ctgtgagcacaagaagccttctctgtgatacgaagtgagatctgtgtgacccaatgatctac 4032  
Qy 3661 tccagagcatcaagaagcaggtgccccgttcagaagttctccagctctacatctcagtt 3720  
Db 4033 tccagagcatcaagaagcaggtgccccgttcagaagttctccagctctacatctcagtt 4092  
Qy 3721 tatcatcatgtatagctcttctgataaggaagtgccagacttatcgaaggaattccactgg 3780  
|||||

Db 4093 tatcatcatgtatagctcttctgataaggaagtggtcagactatccagaaattccactgg 4152  
Qy 3781 aaccttaatgtctctcttggcaatgtggaattcatcgggataaacaataatttttaa 3840  
Db 4153 aaccttaatgtctctcttggcaatgtggaattcatcgggataaacaataatttttaa 4212  
Qy 3841 cccctcaattatttgcctgataacatccgttctgcaacccaacccatctatagcattcgcagcac 3900  
Db 4213 cccctcaattatttgcctgataacatccgttctgcaacccaacccatctatagcattcgcagcac 4272  
Qy 3901 tcttcgcatggaatgtgatagtggtgtgatttaaatagttgcagaatccatctgggaatgga 3960  
Db 4273 tcttcgcatggaatgtgatagtggtgtgatttaaatagttgcagaatccatctgggaatgga 4332  
Qy 3961 gagtaagaacataatcagatcagagatctactgtctacttccatataatgttgc 4020  
Db 4333 gagtaagaacataatcagatcagagatctactgtctacttccatataatgttgc 4392  
Qy 4021 cactgtctccctccaagaagctgcacttcaacctccaagggaggtatgctgagagac 4080  
Db 4393 cactgtctccctccaagaagctgcacttcaacctccaagggaggtatgctgagagac 4452  
Qy 4081 tcaagtgaataatccaaaagatgtgctgcaagtggaacttccaagaagaacatgaaagtccac 4140  
Db 4453 tcaagtgaataatccaaaagatgtgctgcaagtggaacttccaagaagaacatgaaagtccac 4512  
Qy 4141 aggaagttaacttaactcaggaatgaaatctctgtcttaacagcatgtatgtgaagaagttccct 4200  
Db 4513 aggaagttaacttaactcaggaatgaaatctctgtcttaacagcatgtatgtgaagaagttccct 4572  
Qy 4201 cactccagcagtcacaagatgcccacagtggaactcttcttccaagaatgccaagtaaa 4260  
Db 4573 cactccagcagtcacaagatgcccacagtggaactcttcttccaagaatgccaagtaaa 4632  
Qy 4261 ggttttccaaggaatacaagaactcttcaacacctgtgtgtaactcttaagaccacggt 4320  
Db 4633 ggttttccaaggaatacaagaactcttcaacacctgtgtgtaactcttaagaccacggt 4692  
Qy 4321 actgactcgttaccttgaatttcccccagagttgtgggtgacacagaattgcccctgaagat 4380  
Db 4693 actgactcgttaccttgaatttcccccagagttgtgggtgacacagaattgcccctgaagat 4752  
Qy 4381 ggaagttctgtggtcgtgagggcacaggaactctactga 4417  
Db 4753 ggaagttctgtggtcgtgagggcacaggaactctactga 4789  
  
RESULT 9  
AAA49232  
ID AAA49232 standard; DNA; 12445 BP.  
XX  
AC AAA49232;  
XX  
DT 26-SEP-2000 (first entry)  
XX  
DE Vector HSReneo for transforming endothelial cells.  
XX  
XX Hemostatic; antianemic; antidiabetic; nootropic; neuroprotective; vector;  
KW osteopathic; antislaking; immunostimulant; gene therapy; collage;  
KW endothelial cell; peripheral blood; buffy coat cell; VEGF; PCR primer;  
KW vascular endothelial growth factor; bovine brain extract; haemophilia;  
KW Factor VIII; human; transgene; adenovine deaminase deficiency; ss;  
KW sickle cell anemia; thalassemia; diabetes; alpha-antitrypsin deficiency;  
KW Alzheimer's disease; brain disease; heart disease; Immune system defect;  
KW bone fracture; osteoporosis.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
XX WO2000032750-A1.  
XX  
XX 08-JUN-2000.  
XX

PF 24-NOV-1999; 99WC-US28033.  
XX  
PR 24-NOV-1998; 98US-0109687.  
XX  
PA (MINU ) UNIV MINNESOTA.  
PA (UYEM-) UNIV EMORY.  
PA (HEBB/) HEBBEL R P.  
PA (LINY/) LIN Y.  
PA (LOLL/) LOLLAR J S.  
XX  
PI HEBBEL RP, Lin Y, Lollar JS.  
DR WPI: 2000-412303/35.  
XX  
PT Expanding population of endothelial cells useful to biocompatibilize  
PT implantable medical devices comprises contacting buffy coat cells with  
PT collagen I coated surface in culture medium comprising vascular  
PT endothelial growth factor -  
XX  
XX  
PS Claim 19; Fig 4; 53pp; English.  
XX  
XX The invention relates to a method for expanding the population of  
CC endothelial cells (EC) obtained from peripheral blood, by culturing,  
CC in contact with a collagen I coated surface, buffy coat cells obtained  
CC from peripheral mammalian blood in the presence of a culture medium  
CC containing vascular endothelial growth factor (VEGF) and free of bovine  
CC brain extract. EC are useful for treating hemophilia by introducing it  
CC into the blood stream of a mammal, so that an effective amount of  
CC Factor VIII protein is secreted in the blood stream of the mammal.  
CC This sequence represents the vector HSPNeo where the human factor  
CC VIII gene into which an enhanced green fluorescent protein coding  
CC sequence (H8Q) has been inserted, is subcloned. Transgenic EC transduced  
CC in vitro are useful for improving prosthetic implants. EC is also useful  
CC for diagnosing clotting disorders where indication or disease is  
CC associated with a reduction in the activity of an enzyme. EC is also  
CC useful in gene therapy for treating the variety of diseases including  
CC adenosine deaminase deficiency, sickle cell anemia, thalassemia,  
CC hemophilia, diabetes, alpha-antitrypsin deficiency, brain disorders  
CC such as Alzheimer's disease, heart diseases, defects in immune system,  
CC for repairing bone fractures and to treat or prevent osteoporosis.  
XX  
XX  
SQ Sequence 12445 BP; 3255 A; 2989 C; 2951 G; 3250 T; 0 other:  
  
Query Match 97.7%; Score 4315; DB 21; Length 12445;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 4370; Conservative 0; Mismatches 5; Indels 42; Gaps 1;  
  
QY 1 catgcaatagagctccacgtctctctctgtgctcttgagattcgattagtc 60  
DB 767 catgcaatagagctccacgtctctctctgtgctcttgagattcgattagtc 826  
QY 61 caccagaagatctactctgtgtagtgagtgatctcagtgagactatagcaagagatct 120  
DB 827 caccagaagatctactctgtgtagtgagtgatctcagtgagactatagcaagagatct 886  
QY 121 cggctgagctgctgtgagcgaagattctcctcctagagtgccaagattcttccattcaa 180  
DB 887 cggctgagctgctgtgagcgaagattctcctcctagagtgccaagattcttccattcaa 946  
QY 181 caccctcagtcgtgtacaaaagagctcgtttgttagaatccaggttcaactttcaaat 240  
DB 947 caccctcagtcgtgtgtacaaaagagctcgtttgttagaatccaggttcaactttcaaat 1006  
QY 241 cgtctaaagcgaagccacctgagtggtgtctaggtctcaccatccaggtcgaagttta 300  
DB 1007 cgtctaaagcgaagccacctgagtggtgtctaggtctcaccatccaggtcgaagttta 1066  
QY 301 tgaatacagtgatcattcaacttaagaacatgtctccatcctgtcaattcttcagtcgt 360  
DB 1067 tgaatacagtgatcattcaacttaagaacatgtctccatcctgtcaattcttcagtcgt 1126  
QY 361 tgggtatctactcagaaagctcttgaggaagctgaataatgatacagacagtcgaaag 420

DB 1127 tgggtatctactcagaaagctcttgaggaagctgaataatgatacagacagtcgaaag 1186  
QY 421 ggaagaagaagatgataaagctctccctgtgaggaagcacaataatgtctgcaggtcct 480  
DB 1187 ggaagaagaagatgataaagctctccctgtgaggaagcacaataatgtctgcaggtcct 1246  
QY 481 gaaagaagatggtccaaatgagcctctgacccactgtgccttactactactacttcttca 540  
DB 1247 gaaagaagatggtccaaatgagcctctgacccactgtgccttactactactacttcttca 1306  
QY 541 tgtggaacctgtgtaaagagcttgaatcagccctaatgtgagagcctactagatagtaga 600  
DB 1307 tgtggaacctgtgtgtaaagagcttgaatcagccctaatgtgagagcctactagatagtaga 1366  
QY 601 agggagcttgcgcaaggaagaaagacagacacttgcacaaatttactactcttctgtc 660  
DB 1367 agggagcttgcgcaaggaagaaagacagacacttgcacaaatttactactcttctgtc 1426  
QY 661 attgtatgaaggaaagattgtgcaactcagaacaaagactcctgtatgacagataagga 720  
DB 1427 attgtatgaaggaaagattgtgcaactcagaacaaagactcctgtatgacagataagga 1486  
QY 721 tgtgtacatctgtcggcctgtggtcctaaatgacacagtgcaatgtgtatgtaaaggtc 780  
DB 1487 tgtgtacatctgtcggcctgtggtcctaaatgacacagtgcaatgtgtatgtaaaggtc 1546  
QY 781 tctgcgaagctcgtatgtatgacacaggaagaaatcagctcattgtgagattgtgaatgag 840  
DB 1547 tctgcgaagctcgtatgtatgacacaggaagaaatcagctcattgtgagattgtgaatgag 1606  
QY 841 caccactcctgaagtgcacccaataattctcgaagtgacacacattcttctgtgaagaaaca 900  
DB 1607 caccactcctgaagtgcacccaataattctcgaagtgacacacattcttctgtgaagaaaca 1666  
QY 901 tgcgcgaagcgtcctgtggaatctgcgcaataacttcttactgtcctaaacacttgat 960  
DB 1667 tgcgcgaagcgtcctgtggaatctgcgcaataacttcttactgtcctaaacacttgat 1726  
QY 961 ggaacttgcagatcttcacatgattgtgcatatctctccacacaatgatagtcagga 1020  
DB 1727 ggaacttgcagatcttcacatgattgtgcatatctctccacacaatgatagtcagga 1786  
QY 1021 agcttatgtcaaatgtagcagctgtccagaggaacccccaactacgaaatgaaataatga 1080  
DB 1787 agcttatgtcaaatgtagcagctgtccagaggaacccccaactacgaaatgaaataatga 1846  
QY 1081 agaagcggaaagactatgtatgatacttactgattctgaaatgtagtgtgtcaggttga 1140  
DB 1847 agaagcggaaagactatgtatgatacttactgattctgaaatgtagtgtgtcaggttga 1906  
QY 1141 tgaatacaactctcctctcttataccaattcgcctcagttgccaagaagacactctaagc 1200  
DB 1907 tgaatacaactctcctctcttataccaattcgcctcagttgccaagaagacactctaagc 1966  
QY 1201 ttggttatcatcatctgtctgtgaagagagagactgtgactatgtcctcctagtcctgc 1260  
DB 1967 ttggttatcatcatctgtctgtgaagagagagactgtgactatgtcctcctagtcctgc 2026  
QY 1261 ccccgatgacgaaggttataaagtcgaatatttgaacaatgtgcctcagaaggttgtag 1320  
DB 2027 ccccgatgacgaaggttataaagtcgaatatttgaacaatgtgcctcagaaggttgtag 2086  
QY 1321 gaaatgacaataaagtcgaatatttgaatgacatgaaacactttaaagctgtgagac 1380  
DB 2087 gaaatgacaataaagtcgaatatttgaatgacatgaaacactttaaagctgtgagac 2146  
QY 1381 tattcaagatgaatcagaagacttctggaacacttacttattgtggaagtgtggaacacact 1440  
DB 2147 tattcaagatgaatcagaagacttctggaacacttacttattgtggaagtgtggaacacact 2206  
QY 1441 gttgattatatttaagaatcaagcagaagcacaataataacttaacctcaaggaatcac 1500

|    |      |  |      |
|----|------|--|------|
| Dh | 2207 | gttgattatatttaagaatataagaacaagcagaccataataacatctacccctacggatcac  | 2206 |
| Oy | 1501 | tgatgtccgtccctctgttatctcaagsgagaltaccaaaaggtgttaaaaaatttgaagbat  | 1500 |
| Dh | 2267 | tgatgtccgtccctctgttatctcaagsgagaltaccaaaaggtgttaaaaaatttgaagbat  | 2236 |
| Oy | 1501 | tcgaattctgcacagsgaataattcaataataatgagacagtgtactgttaagaagtggcc    | 1620 |
| Dh | 2327 | tcgaattctgcacagsgaataattcaataataatgagacagtgtactgttaagaagtggcc    | 2386 |
| Oy | 1621 | aacaaatcaagatccctccggccttgaccccgctacactacgtattcttcgtttaatggagag  | 1680 |
| Dh | 2387 | aacaaatcaagatccctccggccttgaccccgctacactacgtattcttcgtttaatggagag  | 2446 |
| Oy | 1601 | agatctagcttcagagactcaatltgcccctctccctcactcgtctcaaaaagaatctgtatga | 1740 |
| Dh | 2447 | agatctagcttcagagactcaatltgcccctctccctcactcgtctcaaaaagaatctgtatga | 2506 |
| Oy | 1741 | aagagsgaaccagaataatgttcagaaagaagaatgtcatcctggtttctgtcatattgata   | 1800 |
| Dh | 2507 | aagagsgaaccagaataatgttcagaaagaagaatgtcatcctggtttctgtcatattgata   | 2566 |
| Oy | 1801 | gaaccgagactgtgttacctcaacagagaataataaacgctttcccccataccagctggagt   | 1860 |
| Dh | 2567 | gaaccgagactgtgttacctcaacagagaataataaacgctttcccccataccagctggagt   | 2626 |
| Oy | 1861 | gcagctctgagatccagagattccaaagccctccacaacalcalgtcacagatcaaatgtctat | 1920 |
| Dh | 2627 | gcagctctgagatccagagattccaaagccctccacaacalcalgtcacagatcaaatgtctat | 2686 |
| Oy | 1921 | ttttgatagtttgcagttgttcagtttggtttcgaatgaggtgtgcatactgtcatcttaag   | 1980 |
| Dh | 2687 | ttttgatagtttgcagttgttcagtttggtttcgaatgaggtgtgcatactgtcatcttaag   | 2746 |
| Oy | 1981 | catgtggagacagacagactgcgaactcccttccttccttcctcccttgataataccttcaaaa | 2040 |
| Dh | 2747 | catgtggagacagacagactgcgaactcccttccttccttcctcccttgataataccttcaaaa | 2806 |
| Oy | 2041 | aatgtctatgaagaacacactcaacctatcccatctccagagaagaactgtctcatgtc      | 2100 |
| Dh | 2807 | aatgtctatgaagaacacactcaacctatcccatctccagagaagaactgtctcatgtc      | 2866 |
| Oy | 2101 | gattggaaaacccaggtctctatgtatctctggggtgtccacaactagaacttcggaaagag   | 2160 |
| Dh | 2867 | gattggaaaacccaggtctctatgtatctctggggtgtccacaactagaacttcggaaagag   | 2926 |
| Oy | 2161 | catgaacgcctctctcgtgaagtttctctagttgtgtgaagaagaacatgttgatttttcggag | 2220 |
| Dh | 2927 | catgaacgcctctctcgtgaagtttctctagttgtgtgaagaagaacatgttgatttttcggag | 2986 |
| Oy | 2221 | cagttgaagaagatatttcagacactctgtctgtgttaaaaaatgtccatttgaaccaagag   | 2280 |
| Dh | 2987 | cagttgaagaagatatttcagacactctgtctgtgtgttaaaaaatgtccatttgaacctagag | 3046 |
| Oy | 2281 | ctctcccggaattcaagacacccctagcacatgacaaaagaatttaatgtccacccaccc     | 2340 |
| Dh | 3047 | ctctcccggaattcaagacacccctagcacatgacaaaagaatttaatgtccacccaccc     | 3064 |
| Oy | 2341 | agttcttgaagaacgcatacaacaggaanaataactcgttactactctctcagtaagtaagaga | 2400 |
| Dh | 3065 | agttcttgaagaacgcatacaacaggaanaataactcgttactactctctcagtaagtaagaga | 3124 |
| Oy | 2401 | aattgacataagatcatcaacatcatcatctbaaaatgaagaagagagatttctgacatttga  | 2460 |
| Dh | 3125 | aattgacataagatcatcaacatcatcatctbaaaatgaagaagagagatttctgacatttga  | 3184 |
| Oy | 2461 | tgaagatgaanaatcagagcccccccgagcttccaaaagaagaacagacataatttatgtc    | 2520 |
| Dh | 3185 | tgaagatgaanaatcagagcccccccgagcttccaaaagaagaacacataatttatgtc      | 3244 |
| Oy | 2521 | tgcacgttgagagagctctcggagattatgtgagttgagttccccaatgtttctaaagaanaag | 2580 |
| Dh | 3245 | tgcacgttgagagagctctcggagattatgtgagttgagttccccaatgtttctaaagaanaag | 3304 |

|    |      |  |      |
|----|------|--|------|
| QY | 2581 | ggccagagctggcagctgcccctcaagctcaagaagctggtcttccagggaattctacgagc       | 2640 |
| Db | 3305 | ggcccagagctggcagctgcccctcaagctcaagaagctggtcttccagggaattctacgagc      | 3364 |
| QY | 2641 | ctcccttactcaagcccttaccgctggagagaactcaaatgtaacatttggagctctcgggccc     | 2700 |
| Db | 3365 | ctcccttactcaagcccttaccgctggagagaactcaaatgtaacatttggagctctcgggccc     | 3424 |
| QY | 2701 | atatataagagcagaagcttgaagataataatcatgctgaacttcacgaataatcagcctccg      | 2760 |
| Db | 3425 | atatataagagcagaagcttgaagataataatcatgctgaacttcacgaataatcagcctccg      | 3484 |
| QY | 2761 | tcacctatccctctatcttcagccttctctctctctctctctctctctctctctctctctctct     | 2820 |
| Db | 3485 | tcacctatccctctatcttcagccttctctctctctctctctctctctctctctctctctctct     | 3544 |
| QY | 2821 | accctagaaaaaacttctgccaagcctcaatgnaaccacaacttactcttggaaagtgcaca       | 2880 |
| Db | 3545 | accctagaaaaaacttctgccaagcctcaatgnaaccacaacttactcttggaaagtgcaca       | 3604 |
| QY | 2881 | tcatactggcaccocctcaagaatgaatttgaatgctcaagccttgagcttatctctcgaagt      | 2940 |
| Db | 3605 | tcatactggcaccocctcaagaatgaatttgaatgctcaagccttgagcttatctctcgaagt      | 3664 |
| QY | 2941 | tgaccttgnaaaaagatgtgcaactcagaagcctgtaattgnaaccctctctgtctgcacactaa    | 3000 |
| Db | 3665 | tgaccttgnaaaaagatgtgcaactcagaagcctgtaattgnaaccctctctgtctgcacactaa    | 3724 |
| QY | 3001 | cacacgtgaacccctgtcctcatgaggagacaagtgtgaactgacagaattgtctctgtttttac    | 3060 |
| Db | 3725 | cacacgtgaacccctgtcctcatgaggagacaagtgtgaactgacagaattgtctctgtttttac    | 3784 |
| QY | 3061 | catctttgatatgagcccaagaagctgctgaacttcaactgtgaataatgagaaagaactgtgaagcc | 3120 |
| Db | 3785 | catctttgatatgagcccaagaagctgctgaacttcaactgtgaataatgagaaagaactgtgaagcc | 3844 |
| QY | 3121 | tcacctgcaatatccagaatggaagatcccaactttaagaagaattatcgcttccatgtcaat      | 3180 |
| Db | 3845 | tcacctgcaatatccagaatggaagatcccaactttaagaagaattatcgcttccatgtcaat      | 3904 |
| QY | 3181 | caatgctgtaacataatgataacacttactctggttaagaaagcttccagatcaaaagattcg      | 3240 |
| Db | 3905 | caatgctgtaacataatgataacacttactctggttaagaaagcttccagatcaaaagattcg      | 3964 |
| QY | 3241 | atggtatctgctcagcagctggcagcaatgnaaacatcatcttatcttattcagtgagca         | 4024 |
| Db | 3965 | atggtatctgctcagcagctggcagcaatgnaaacatcatcttatcttattcagtgagca         |      |
| QY | 3301 | tgtgttcaactgttaccgaaaaaagaagatataaatgtgcactgttacaattcttaccag         | 3360 |
| Db | 4025 | tgtgttcaactgttaccgaaaaaagaagatataaatgtgcactgttacaattcttaccag         | 4084 |
| QY | 3361 | tgtttttgagacagctggaataatgttaccatcccaagcttgtaatttgcgcggttggaattgct    | 3420 |
| Db | 4085 | tgtttttgagacagctggaataatgttaccatcccaagcttgtaatttgcgcggttggaattgct    | 4144 |
| QY | 3421 | tatttgcgagatcatataatgctgtggaatgagacacttttctgtgtgtacagaacaataagtg     | 3480 |
| Db | 4145 | tatttgcgagatcatataatgctgtggaatgagacacttttctgtgtgtgtacagaacaataagtg   | 4204 |
| QY | 3481 | tcgaactccoccttggaattgctctctggaacatctgaagaatttccagatataacgcttcaag     | 3540 |
| Db | 4205 | tcgaactccoccttggaattgctctctggaacatctgaagaatttccagatataacgcttcaag     | 4264 |
| QY | 3541 | acaatattggaacagtgtggcccccaagaagcttggccagacttctatattcccgatcaaatatgc   | 3600 |
| Db | 4265 | acaatattggaacagtgtggcccccaagaagcttggccagacttctatattcccgatcaaatatgc   | 4324 |
| QY | 3601 | cttggagacccaagaagcccttctcttggatccaagtggtgactgttggcaccacaatgat        | 3660 |
| Db | 4325 | cttggagacccaagaagcccttctcttggatccaagtggtgactgttggcaccacaatgat        | 4384 |

|           |  |  |      |
|-----------|--|--|------|
| OY        | 3661   | taagcagcatcaaaagcccaagagtgcccgtaagaattcttccaagcctctacatctcgaatt    | 3720 |
| Db        | 4385   | tcaaggcatcaaaagcccaagagtgcccgtaagaattcttccaagcctctacatctcgaatt     | 4444 |
| OY        | 3721   | tatcatcatgtataagctcttgaatggagaagatggcagactatcgaagaaattccacg        | 3780 |
| Db        | 4445   | tatcatcatgtataagctcttgaatggagaagatggcagactatcgaagaaattccacg        | 4504 |
| OY        | 3781   | aaccttaatggtctctcttggcaaatgtgagatcatctgggaataaacaataatttttaa       | 3840 |
| Db        | 4505   | aaccttaatggtctctcttggcaaatgtgagatcatctgggaataaacaataatttttaa       | 4564 |
| OY        | 3841   | ccctccaattattgtctcgtatgatacatcggtttgaccaccaattatataagatctggaagac   | 3900 |
| Db        | 4565   | ccctccaattattgtctcgtatgatacatcggtttgaccaccaattatataagatctggaagac   | 4624 |
| OY        | 3901   | tcttcgcataagatctgaatgggctctgtatcttaaatagtgtgcagcatggcatctggagaatga | 3960 |
| Db        | 4625   | tcttcgcataagatctgaatgggctctgtatcttaaatagtgtgcagcatggcatctggagaatga | 4684 |
| OY        | 3961   | gagtaagaagcaatactcagatgcacagaattactgcttcatctcacttlaacaaatagtttgc   | 4020 |
| Db        | 4685   | gagtaagaagcaatactcagatgcacagaattactgcttcatctcacttlaacaaatagtttgc   | 4744 |
| OY        | 4021   | caactggtctctccttcaaaagctcgacttcaacctccaagggagagtaatgccttggagacc    | 4080 |
| Db        | 4745   | caactggtctctccttcaaaagctcgacttcaacctccaagggagagtaatgccttggagacc    | 4804 |
| OY        | 4081   | tcaaggtaataatccaaagaagtgtgcgcagaatggaattccagaagacataatgaagtac      | 4140 |
| Db        | 4805   | tcaaggtaataatccaaagaagtgtgcgcagaatggaattccagaagacataatgaagtac      | 4864 |
| OY        | 4141   | aggagtaactactccaggaggatlaaatctctgcttaccagcatgtatagtgaaggagttcct    | 4200 |
| Db        | 4865   | aggagtaactactccaggaggatlaaatctctgcttaccagcatgtatagtgaaggagttcct    | 4924 |
| OY        | 4201   | catctccaagcatgaagaatggccatcagttgactctcttttttcaagaatggcaagaataa     | 4260 |
| Db        | 4925   | catctccaagcatgaagaatggccatcagttgactctcttttttcaagaatggcaagaataa     | 4984 |
| OY        | 4261   | gatttttcagggaataatcaagaactctctcaacacctgtgtgtaacctctagaaccacggt     | 4320 |
| Db        | 4985   | gatttttcagggaataatcaagaactctctcaacacctgtgtgtaacctctagaaccacggt     | 5044 |
| OY        | 4321   | actgactcgtactacttcaaatcaacccccaagatttgggtgcacagatctgccttgaagat     | 4380 |
| Db        | 5045   | actgactcgtactacttcaaatcaacccccaagatttgggtgcacagatctgccttgaagat     | 5104 |
| OY        | 4381   | ggaggttcttggtgctcgagagcacagaagcctctactga                           | 4417 |
| Db        | 5105   | ggaggttcttggtgctcgagagcacagaagcctctactga                           | 5141 |
| result_10 |  |  |      |
| AAD00121  |  |  |      |
| ID        | AAD00121 standard; DNA; 11933 BP.  |  |      |
| XX        | AAD00121;  |  |      |
| XX        | 31-JUL-2000 (first entry)  |  |      |
| DE        | Recombinant adeno associated vector construct, PAAV-F8-1.                |  |      |
| KW        | Recombinant Adeno Associated Vector: rAAV: PAAV-F8-1; human Factor VIII; |  |      |
| KW        | hFVIII; hNF-3 albumin promoter; human elongation factor-1alpha; E1alpha  |  |      |
| KW        | human growth hormone; hGH; inverted terminal repeat; ITR; haemophilia;   |  |      |
| XX        | gene therapy; ds.  |  |      |
| OS        | Adeno associated virus.  |  |      |
| PN        | WO200023116-A1.  |  |      |
| PD        | 27-APR-2000.   |  |      |

|    |  |   |               |
|----|--|---|---------------|
| XX | PF   | 19-OCT-1999;  | 99NC-US24495. |
| XX | PR   | 20-OCT-1998;  | 98US-0104994. |
| XX | PR   | 24-MAR-1999;  | 99US-0125974. |
| XX | PR   | 30-JUL-1999;  | 99US-0364862. |
| XX | PA   | (AVIG-) AVIGEN INC.   |               |
| XX | PI   | Couto LB, Colosi PC;  |               |
| XX | DR   | WPI: 2000-339536/29.  |               |
| XX | PT   | New recombinant adenovirus-associated vector, useful for gene therapy |               |
| XX | PT   | to treat hemophilia, comprises at least a portion of Factor VIII      |               |
| XX | PT   | operably linked to control sequence                                   |               |
| PS | Example 2; Fig 5; 92pp; English.   |   |               |
| CC | The present DNA sequence is a recombinant adenovirus-associated vector,    |   |               |
| CC | (AAV) construct, pAAV-F8-1. This expression vector comprises the HM-3      |   |               |
| CC | albumin promoter, the first intron (-573 to +985) of human elongation      |   |               |
| CC | factor-1alpha (EFL1alpha) gene, human Factor VIII coding sequence (hFVIII) |   |               |
| CC | and polyadenylation signal from human growth hormone (hGH). This sequence  |   |               |
| CC | is inserted between the AAV inverted terminal repeat (ITR) regions. The    |   |               |
| CC | hFVIII coding region comprises the heavy chain gene segment with the       |   |               |
| CC | first 57 bp encoding the 19 amino acid signal peptide, A1 and A2 domains   |   |               |
| CC | and 5 amino acids from the N-terminus of the B domain. The light chain     |   |               |
| CC | segment comprises the C-terminal 85 amino acids of B domain and the A3,    |   |               |
| CC | C1 and C2 domains. Both the heavy and light chain segments are cloned      |   |               |
| CC | into the same plasmid separated by 42 nucleotides coding for 14 residues   |   |               |
| CC | of the B domain, that is deleted. This plasmid is operably linked to       |   |               |
| CC | control sequences, that directs the transcription and translation of the   |   |               |
| CC | Factor VIII gene. The adeno-associated viral vectors are used for gene     |   |               |
| CC | therapy to treat haemophilia. This method allows prolonged expression of   |   |               |
| CC | therapeutic levels of Factor VIII in vivo. The rAAV are used for gene      |   |               |
| CC | therapy, because of their broad host range, safety profile and duration    |   |               |
| CC | of expression in the infected hosts.                                       |   |               |
| SQ | Sequence 11933 BP; 3258 A; 2818 C; 2717 G; 3140 T; 0 other:                |   |               |
| QY | Query Match  | 96.2%; Score 4249.4; DB 21; Length 11933;                             |               |
| QY | Best Local Similarity  | 98.2%; Pred. No. 0;   |               |
| QY | Matches 4338; Conservative   | 0; Mismatches 1; Indels 78; Gaps 1                                    |               |
| DB | 1 catcaaatagagtccacccgcttccttcttgtagcattggagatctgttaagtc                   | 60  |               |
| DB | 426 catgcacaataagaagctcccaccgcttcttcttgtagcattggagatctgttaagtc             | 485   |               |
| OY | 61 caccagaagatactacctggtgtgacltggaactgtcatatggaatatgatcaaaigtatct          | 120   |               |
| DB | 486 caccagaagatactacctggtgtgacltggaactgtcatatggaatatgatcaaaigtatct         | 545   |               |
| OY | 121 cggtgagctgcctgttgagcgcaagatttccctctagagtgccaatattttcatctcaa            | 180   |               |
| DB | 546 cggtgagctgcctgttgagcgcaagatttccctctagagtgccaatattttcatctcaa            | 605   |               |
| OY | 181 caactcagctgttgacaanaaagacttgtttttagaatccaaggttcaacctttacaacat          | 240   |               |
| DB | 606 caactcagctgttgacaanaaagacttgtttttagaatccaaggttcaacctttacaacat          | 665   |               |
| OY | 241 cgctaaagccaagggcacccctggacltggtctgtctaggctccactaacacagagctgagtta       | 300   |               |
| DB | 666 cgctaaagccaagggcacccctggacltggtctgtctaggctccactaacacagagctgagtta       | 725   |               |
| OY | 301 tgatacagtggtcatctaacttaagaacatlggtctccactccctgcagttcttaagtct           | 360   |               |
| DB | 726 tgatacagtggtcatctaacttaagaacatlggtctccactccctgcagttcttaagtct           | 785   |               |
| OY | 361 tgggttatctactatcgaaaagcttcctgaagggagctgaatatgatgatcaagaccagtaaa        | 420   |               |
| DB | 786 tgggttatctactatcgaaaagcttcctgaagggagctgaatatgatgatcaagaccagtaaa        | 845   |               |



QY 421 ggaagaaagatgatcaaaagtcttccttgg tggaaagccatacatatgtctggcaggtcct 480  
|||||  
Db 846 gggagaaagaaagatgataaagctcttccttgg tgaagccatacatatgtctggcaggtcct 905  
QY 481 gaaagaaatggtcccaatggtcctctga ccccaactgtgtccttactactacataltcttctca 540  
|||||  
Db 906 gaaagaaatggtcccaatggtcctctga ccccaactgtgtccttactactacataltcttctca 965  
QY 541 tgtggagcgtgtataaagacttgaattcag gctccatttggagccctactagtatgttaga 600  
|||||  
Db 966 tgtggaccttgtataaagacttgaattcag gctccatttggagccctactagtatgttaga 1025  
QY 601 agggactctggcacaagaaagacacagac cactgtgcacaatatattactacttcttctgt 660  
|||||  
Db 1026 agggactctggcacaagaaagacacagac cactgtgcacaatatattactacttcttctgt 1085  
QY 661 atttgaatgaagggaagaaatgtgcactcag aagaaactccttgaatgcagagataggga 720  
|||||  
Db 1086 atttgaatgaagggaagaaatgtgcactcag aagaaactccttgaatgcagagataggga 1145  
QY 721 tgtctcatctgtctgggcttggccttaaat gcaacagttcaatggttatagttaacaggtc 780  
|||||  
Db 1146 tgtctcatctgtctgggcttggccttaaat gcaacagttcaatggttatagttataacaggtc 1205  
QY 781 tctggcaggtcttattgattggatgcacaga ggaatcagttcttggcagtgatgtgaattgg 840  
|||||  
Db 1206 tctggcaggtcttattgattggatgcacaga ggaatcagttcttggcagtgatgtgaattgg 1265  
QY 841 caccactctctgaagtgactcaataatctccg aaggtcacaacatcttctgttagagacca 900  
|||||  
Db 1266 caccactctctgaagtgactcaataatctccg aaggtcacaacatcttctgttagagacca 1325  
QY 901 tggcagagcgtctctgtgaatactcgcgaata acttcccttactgtctcaaacctcttgat 960  
|||||  
Db 1326 tggcagagcgtctctgtgaatactcgcgaata acttcccttactgtctcaaacctcttgat 1385  
QY 961 gggacttggacagatttctactgttcttgta ctaactcttccacacaatcagatggacatga 1020  
|||||  
Db 1386 gggacttggacagatttctactgttcttgta ctaactcttccacacaatcagatggacatga 1445  
QY 1021 agcttattgtcaaaatgtagacagctgttcc aagagaaacccaactaagatlgaaaaataatga 1080  
|||||  
Db 1446 agcttattgtcaaaatgtagacagctgttcc aagagaaacccaactaagatlgaaaaataatga 1505  
QY 1081 agaaagcggaaagatatagtatgatacttact gatacttctgaaatggatgtggtcaggtttga 1140  
|||||  
Db 1506 agaaagcggaaagatatagtatgatacttact gatacttctgaaatggatgtggtcaggtttga 1565  
QY 1141 tgaatgaacactcctccttcttacttcaaat ttcgctcagttggccaagaagacatcttaaaac 1200  
|||||  
Db 1566 tgaatgaacactcctccttcttacttcaaat ttcgctcagttggccaagaagacatcttaaaac 1625  
QY 1201 ttgggtacattacatgtctgtctgaagaggaa gactgtggaactatgtctccctttagtctcgc 1260  
|||||  
Db 1626 ttgggtacattacatgtctgtctgaagaggaa gactgtggaactatgtctccctttagtctcgc 1685  
QY 1261 ccccgatgacagaagtctataaagaatccaata ttgaaacatgtgcctcagcggatgtgtag 1320  
|||||  
Db 1686 ccccgatgacagaagtctataaagaatccaata ttgaaacatgtgcctcagcggatgtgttag 1745  
QY 1321 gaagtaacaaaagtcggactttagcatatacaga taaagtaaacctttaaagactcgtgaagc 1380  
|||||  
Db 1746 gaagtaacaaaagtcggactttagcatatacaga taaagtaaacctttaaagactcgtgaagc 1805  
QY 1381 tattcagcatlgaaatcaggaatcttgggaccttact ttaatttggggaagtgtggagaccact 1440  
|||||  
Db 1806 tattcagcatlgaaatcaggaatcttgggaccttact ttaatttggggaagtgtggagaccact 1865  
QY 1441 gttgattatatttaagaatccaagcaagacga ccatataacatctaccctcaeggaattcac 1500  
|||||  
Db 1866 gttgattatatttaagaatccaagcaagacga ccatataacatctaccctcaeggaattcac 1925

QY 1501 tgaatgcgtccttcttattcaagggagattacc aaaaagtgtaaaaacatttgaagattt 1560  
|||||  
Db 1926 tgaatgcgtccttcttattcaagggagattacc aaaaagtgtaaaaacatttgaagattt 1985  
QY 1561 tccaattctgcagggagaaatattccaataataa tggacagtgactgtagaagaatgggcc 1620  
|||||  
Db 1986 tccaattctgcagggagaaatattccaataataa tggacagtgactgtagaagaatgggcc 2045  
QY 1621 aactaaatcagatccctcgggtgcttgaacggc ttaattactcagtttggttaatatgtgag 1680  
|||||  
Db 2046 aactaaatcagatccctcgggtgcttgaacggc ttaattactcagtttggttaatatgtgag 2105  
QY 1681 agatcttagcttcagagactcaatgtggccctcc tccatctgtcttaaaaagaatctgtatca 1740  
|||||  
Db 2106 agatcttagcttcagagactcaatgtggccctcc tccatctgtcttaaaaagaatctgtatca 2165  
QY 1741 aagaggaaaaccaataatgttcaagaagaagaa tgtcaatctgttcttctgtattttatga 1800  
|||||  
Db 2166 aagaggaaaaccaataatgttcaagaagaagaa tgtcaatctgttcttctgtattttatga 2225  
QY 1801 gaacccgaagctgtgttacctcaagagaataa cagcgtcttctcccaatccagctggagt 1860  
|||||  
Db 2226 gaacccgaagctgtgttacctcaagagaataa cagcgtcttctcccaatccagctggagt 2285  
QY 1861 gcaagcttgaagatccagaagtttccaagcctt ccaacatcatgtacagacataatgtgtatgt 1920  
|||||  
Db 2286 gcaagcttgaagatccagaagtttccaagcctt ccaacatcatgtacagacataatgtgtatgt 2245  
QY 1921 tttctgaatgttctgaatgtgttcaatgtgttct gtaacatggatgtacatgttctctgaag 1980  
|||||  
Db 2346 tttctgaatgttctgaatgtgttcaatgtgttct gtaacatggatgtacatgttctctgaag 2405  
QY 1981 catttgaagcacaagactgactccttctgtctc tctctctctctctgtgatatcaacttcaaca 2040  
|||||  
Db 2406 catttgaagcacaagactgactccttctgtctc tctctctctctctgtgatatcaacttcaaca 2465  
QY 2041 aatgtgtctaatgaagacaacacttccatcccat tctccagggagaaactgtcttccatgtc 2100  
|||||  
Db 2466 aatgtgtctaatgaagacaacacttccatcccat tctccagggagaaactgtcttccatgtc 2525  
QY 2101 gatggaaaacccaagttctatlgatatctctg ggtgcacaacactcagacttctcgaaacag 2160  
|||||  
Db 2526 gatggaaaacccaagttctatlgatatctctg ggtgcacaacactcagacttctcgaaacag 2585  
QY 2161 catgaccgccttactcgtaaagtttctagttgt tgaacaagaacagcgtgtattatacagga 2220  
|||||  
Db 2586 catgaccgccttactcgtaaagtttctagttgt tgaacaagaacagcgtgtattatacagga 2645  
QY 2221 cagttatgaagatatttcaagcatactgtcgtga gtaaaaacaatgycatcttgaaccaagaag 2280  
|||||  
Db 2646 cagttatgaagatatttcaagcatactgtcgtga gtaaaaacaatgycatcttgaaccaagaag 2705  
QY 2281 ctcttcccgaaattcaagacaaccttagcac taaagcaaaatlttaattgccaccacac 2340  
|||||  
Db 2706 ctcttcccgaaattcaagacaaccttagcac taaagcaaaatlttaattgccaccacac 2709  
QY 2341 agtcttggaaaacgccatcaacgggaataact cgttactactcttcagttccagtaacagga 2400  
|||||  
Db 2710 -----gaaaataactcgttactactcttcag ttcagttccagtaacagga 2747  
QY 2401 aattgactatgatgataccatatactcagttga aatgaagaagaagatttggacatttata 2460  
|||||  
Db 2748 aattgactatgatgataccatatactcagttga aatgaagaagaagatttggacatttata 2807  
QY 2461 tgaagatgaanaatcagagcccccgcagcttcca aagaagaacacgcacactattttatgc 2520  
|||||  
Db 2808 tgaagatgaanaatcagagcccccgcagcttcca aagaagaacacgcacactattttatgc 2867  
QY 2521 tgcagtggaagagcctctgggattatagggatag ttagtctcccaatggttcttaagaacag 2580  
|||||  
Db 2868 tgcagtggaagagcctctgggattatagggatag ttagtctcccaatggttcttaagaacag 2927  
QY 2581 ggtctcaagtggtgcaggtgtccctcagttcca agaaagtgttttccaggaatttactgtag 2640





XX 09-SEP-1988; 88EP-0114769.  
PF  
XX  
XX 08-APR-1988; 88JP-0085454.  
PR  
XX  
XX (KAGA ) CHEMO-SERO-THERAP.  
PA (TEIJ ) TEIJIN LTD.  
XX  
XX Sugiyama T, Masuda K, Tajima Y, Yonemura H;  
PI  
PI WPI; 1989-078467/11.  
DR  
DR P-PSDB; AAP91165.  
XX  
XX Prodn. of recombinant human Factor-VIII-C -  
PI using animal cells transformed with a vector contg. the gene for  
PI Factor VIII:C and a promoter  
XX  
XX Fig 1(1) - 1(13); : 32pp; English.  
PS  
XX  
XX When translated, Arg-740 of the carboxyl terminus of the H chain is  
CC directly bonded by a peptide bond to Glu-1649 of the amino terminus of  
CC L chain. It is used to transform animal cells so that they produce  
CC human Factor VIII:C. A prefd. expression vector is plasmid Ad.RE.neo.  
CC The expression vector has at least one promoter upstream of AAN90654.  
CC The transformants can constantly and continuously produce human Factor  
CC VIII:C in high yield on a commercial scale. The human Factor VIII:C so  
CC produced is considered to corresp. to the smallest species of active and  
CC intact Factor VIII:C molecules in the human blood plasma. It is useful  
CC for treating haemophilia A patients.  
XX  
XX Sequence 4275 BP; 1245 A; 941 C; 945 G; 1144 T; 0 other;

| Query Match           | 94.5%;   | Score 4173;  | DB 10;        | Length 4275; |
|-----------------------|--|--------------|---------------|--------------|
| Best Local Similarity | 96.0%;   | Prod. No. 0; |               |              |
| Matches 4270;         | Conservative   | 0;           | Mismatches 5; | Indels 84;   |
|                       |  |              |               | Gaps 1;      |
| QY. 59                | gccaccagaagatctactctggtgcagtgagacgtcattggagactatgatgaaagtat          | 118          |               |              |
| Db 1                  | gccaccagaagatctactctggtgcagtgagacgtcattggagactatgatgaaagtat          | 60           |               |              |
| QY 119                | ctcgtgtgagctcctcgttggacgcagaagattccctccctaagtcgccaatctttccatc        | 178          |               |              |
| Db 61                 | ctcgtgtgagctcctcgttggacgcagaagattccctccctaagtcgccaatctttccatc        | 120          |               |              |
| QY 179                | aacacctgaatcgtgtgtacaaaaagactctgtttgtagaattcaaggttcaacctttcaac       | 238          |               |              |
| Db 121                | aacacctgaatcgtgtgtacaaaaagactctgtttgtagaattcaaggttcaacctttcaac       | 180          |               |              |
| QY 239                | atcgcgtaaagccaaagccaccctcgatgtggtctgcctgaagtcctaccatccagctgaagtt     | 298          |               |              |
| Db 161                | atcgcgtaaagccaaagccaccctcgatgtggtctgcctgaagtcctaccatccagctgaagtt     | 240          |               |              |
| QY 299                | tatgataaagtgtgtcatctatctaaccttaagaacatgtgcttccacatccgttcatgtatgct    | 358          |               |              |
| Db 241                | tatgataaagtgtgtcatctatctaaccttaagaacatgtgcttccacatccgttcatgtatgct    | 300          |               |              |
| QY 359                | gttgtgtatctactatgtgaaagctctctgaaggagctctgatatatgataatcagaaccagttcaa  | 418          |               |              |
| Db 301                | gttgtgtatctactatgtgaaagctctctgaaggagctctgatatatgataatcagaaccagttcaa  | 360          |               |              |
| QY 419                | aggaggaagaagaagatgatcaaaagtcttccctctggtggaagccatacatatgtctcggaagtc   | 478          |               |              |
| Db 361                | aggaggaagaagaagatgatcaaaagtcttccctctggtggaagccatacatatgtctcggaagtc   | 420          |               |              |
| QY 479                | ctgaaagaagaatgtgtccaaatgtgacctctgacccaactgtgcttaccatacatatcttct      | 538          |               |              |
| Db 421                | ctgaaagaagaatgtgtccaaatgtgacctctgacccaactgtgcttaccatacatatcttct      | 480          |               |              |
| QY 539                | catgtgagacccgtgtataaagacttgaattcagaagctccatttgagaagccctactgatatgtaga | 598          |               |              |
| Db 481                | catgtgagacccgtgtataaagacttgaattcagaagctccatttgagaagccctactgatatgtaga | 540          |               |              |

|    |      |  |      |
|----|------|--|------|
| QY | 539  | gaagagatctgcccagaagaaagacacagacctgycacaattatctactcttgcct           | 658  |
| Db | 541  | gaaggaatctcgcgaagagaagaacacacacctctgcacaaatttactactcttgcct         | 600  |
| QY | 659  | gattttgtgaagggaagaagtgtgcctcgaagaacaaagaaacctctgtgtgaagatag        | 718  |
| Db | 601  | gattctgtatgaagggaagaagtgtgcctcgaagaacaaagaaacctctgtatgaagatag      | 660  |
| QY | 719  | gatgtctgacatctgcagggccctgcctcaaaatgcacacgctcaatgtgtatgtlaaacgg     | 778  |
| Db | 661  | gatgtctgacatctgcctgcagggccctgcctcaaaatgcacacgctcaatgtgtatgtlaaacgg | 720  |
| QY | 779  | tctctgcgaagctctgatgtgaatgcccacaggaatacagctcatatgcatgtgtatgtgaatg   | 838  |
| Db | 721  | tctctgacaaagctctgatgtgtgtgcacaggaatacagctcatatgcatgtgtatgtgaatg    | 780  |
| QY | 839  | ggcacacctctctgaagtgacatcaataatctctgtgaaggtgcacacattctctgtgaagac    | 898  |
| Db | 781  | ggcacacctctctgaagtgacatcaataatctctgtgaaggtgcacacattctctgtgaagac    | 840  |
| QY | 899  | catgcgcagcgctcctctggaataatctgcacataactctctactctctcaaacactctg       | 958  |
| Db | 841  | catgcgcagcggtcctctggaataatctgcacataactctctactctctcaaacactctg       | 900  |
| QY | 959  | atggaaccttgagacgtttctactactgtttgttcatactctctccacaacatgcatgtgcatg   | 1018 |
| Db | 901  | atggaaccttgagacgtttctactactgtttgttcatactctctccacaacatgcatgtgcatg   | 960  |
| QY | 1019 | gaagtttatgtccaagtagaagcagctgtgcaggaagaaacccaactaogaaatgaanaatat    | 1078 |
| Db | 961  | gaagtttatgtccaagtagaagcagctgtgcaggaagaaacccaactaogaaatgaanaatat    | 1020 |
| QY | 1079 | gaagaagcgggaagacataatgatagatcttaacgattccgaaatgtgatgtgtgcagttt      | 1138 |
| Db | 1021 | gaagaagcgggaagacataatgatagatcttaacgattccgaaatgtgatgtgtgtgcagttt    | 1080 |
| QY | 1139 | gatgatagaacaactctccctcccttatccaaatctgcctcagttgcaggaagaacatccataa   | 1198 |
| Db | 1081 | gatgatagaacaactctccctcccttatccaaatctgcctcagttgcaggaagaacatccataa   | 1140 |
| QY | 1199 | acttggtgaacattacatctgtctgcctgaaggaagagagactggtgactatgctctcttaactc  | 1258 |
| Db | 1141 | acttggtgaacattacatctgtctgcctgaaggaagagagactggtgactatgctctcttaactc  | 1200 |
| QY | 1259 | gcccccgatgcagaagtgttaaaagcgaattgtgaacaatgtgcctcaaggcgagttgt        | 1318 |
| Db | 1201 | gcccccgatgcagaagtgttaaaagcgaattgtgaacaatgtgcctcaaggcgagttgt        | 1260 |
| QY | 1319 | aggaagtaacaaaagctcgcgatttatgtgcatacacagatgtaaaccttlaagactctgttga   | 1378 |
| Db | 1261 | aggaagtaacaaaagctcgcgatttatgtgcatacacagatgtaaaccttlaagactctgttga   | 1320 |
| QY | 1379 | gctattcagcatgtgaatcaaggaatccttgggaaccttactttaagggaagtgtggaacaca    | 1438 |
| Db | 1321 | gctattcagcatgtgaatcaaggaatccttgggaaccttactttaagggaagtgtggaacaca    | 1380 |
| QY | 1439 | ctgtgtgataattatgaagaaatcaagcaagcagacataaacaatctacccttcaggaatc      | 1498 |
| Db | 1381 | ctgtgtgataattatgaagaaatcaagcaagcagacataaacaatctacccttcaggaatc      | 1440 |
| QY | 1499 | actgatytcctgtctctctgtataatcaaggaatatccacaaaggtgtataaacatttgaagat   | 1558 |
| Db | 1441 | actgatytcctgtctctctgtataatcaaggaatatccacaaaggtgtataaacatttgaagat   | 1500 |
| QY | 1559 | tttccaattctgcaggaagaataatccaataataatgtgacagtgactcgttagaagaatgg     | 1618 |
| Db | 1501 | tttccaattctgcaggaagaataatccaataataatgtgacagtgactcgttagaagaatgg     | 1560 |
| QY | 1619 | ccaactaaatcaagctctgcgtgtgcgcgaacccggtatctactcagtttcgtttaaataagg    | 1678 |
| Db | 1561 | ccaactaaatcaagctctgcgtgtgcgcgaacccggtatctactcagtttcgtttaaataagg    | 1620 |
| QY | 1679 | agaatctagcttcaagcaccatctgtgcctctccatctatctgtataaagaatctgtaat       | 1738 |





|    |      |  |      |
|----|------|--|------|
| OY | 899  | catctgcagagcygtcctcttggaatctcgcgaatacttctcctaactgtctcaacaactctg              | 958  |
| Db | 844  | catctgcagagcygtcctcttggaatctcgcgaatacttctcctaactgtctcaacaactctg              | 903  |
| OY | 959  | atggaccttggaagagtttcttcaactggttttgcacatactcttcccaacaacttgatgcatg             | 1018 |
| Db | 904  | atggaccttggaagagtttcttcaactggttttgcacatactcttcccaacaacttgatgcatg             | 963  |
| OY | 1019 | gaagactatgtccaagtaagtcagacgtctccagaggaaccccaactacgaatagaataataat             | 1078 |
| Db | 964  | gaagactatgtccaagtaagtcagacgtctccagaggaaccccaactacgaatagaataataat             | 1023 |
| OY | 1079 | gaagaagcgggaagactatgatagtatcctaactgattctgaaatbga.tgtgttcaggttctt             | 1138 |
| Db | 1024 | gaagaagcgggaagactatgatagtatcctaactgattctgaaatbga.tgtgttcaggttctt             | 1083 |
| OY | 1139 | gattgatgacaacctctctctcttatacccaattcgcgtcagtttgcgaagaagcatccctaaa             | 1198 |
| Db | 1084 | gattgatgacaacctctctctcttatacccaattcgcgtcagtttgcgaagaagcatccctaaa             | 1143 |
| OY | 1199 | actctggatcatcatcatctgcctctggaagaagagagacctgagactatgtctcccttaagtcctc          | 1258 |
| Db | 1144 | actctggatcatcatcatctgcctctggaagaagagagacctgagactatgtctcccttaagtcctc          | 1203 |
| OY | 1259 | gcccccgatgacagaagattataaagaatcaatatttgacaatgagccctcagcggatgtgtc              | 1318 |
| Db | 1204 | gcccccgatgacagaagattataaagaatcaatatttgacaatgagccctcagcggatgtgtc              | 1263 |
| OY | 1319 | aggaagatcaaaaaaagtccgattatgtggcatacaaga.tgaaaacctttaagacctcggtgaa            | 1378 |
| Db | 1264 | aggaagatcaaaaaaagtccgattatgtggcatacaaga.tgaaaacctttaagacctcggtgaa            | 1323 |
| OY | 1379 | gctatctcagcatatgatactcaagaaactcttggaccttacttata.tgggggaagtctggagacaca        | 1438 |
| Db | 1324 | gctatctcagcatatgatactcaagaaactcttggaccttacttata.tgggggaagtctggagacaca        | 1383 |
| OY | 1439 | ctgttgatctatatatttaagaatccaagcagaagacatataactctatacctccctacggagatc           | 1498 |
| Db | 1384 | ctgttgatctatatatttaagaatccaagcagaagacatataactctatacctccctacggagatc           | 1443 |
| OY | 1499 | actgatgctccgtccctctgtataccaagaga.tttccaaagaagtgtaaacaactttgaagat             | 1558 |
| Db | 1444 | actgatgctccgtccctctgtataccaagaga.tttccaaagaagtgtaaacaactttgaagat             | 1503 |
| OY | 1559 | tttccaactctgcagagagaataatccaataataaat.tgscag.tgactgtagaagaatgagg             | 1618 |
| Db | 1504 | tttccaactctgcagagagaataatccaataataaat.tgscag.tgactgtagaagaatgagg             | 1563 |
| OY | 1619 | ccaactcaaatcaga.tctctcgtgtgccttgagacccgcctataactctg.tttcgttaata.tgag         | 1678 |
| Db | 1564 | ccaactcaaatcaga.tctctcgtgtgccttgagacccgcctataactctg.tttcgttaata.tgag         | 1623 |
| OY | 1679 | agagatctagcttcagagacatcttggccctctccctca.tctgctcaagaagaatctgtatg              | 1738 |
| Db | 1624 | agagatctagcttcagagacatcttggccctctccctca.tctgctcaagaagaatctgtatg              | 1683 |
| OY | 1739 | caaaagcggaaaccagataat.tg.tgcagacaagaaga.tg.tca.tcc.tgt.tttctgtata.ttgat      | 1798 |
| Db | 1684 | caaaagcggaaaccagataat.tg.tgcagacaagaaga.tg.tca.tcc.tgt.tttctgtata.ttgat      | 1743 |
| OY | 1799 | gagaaacaggagctg.tg.tac.tccccaagagaata.tatbcaaaagct.tttccccaatccagctgga       | 1858 |
| Db | 1744 | gagaaacaggagctg.tg.tac.tccccaagagaata.tatbcaaaagct.tttccccaatccagctgga       | 1803 |
| OY | 1859 | gtgcagcttgaagatccaga.tg.tcccaagctccaca.tca.tg.tccagatcatcatg.tgcat           | 1918 |
| Db | 1804 | gtgcagcttgaagatccaga.tg.tcccaagctccaca.tca.tg.tccagatcatcatg.tgcat           | 1863 |
| OY | 1919 | gtttcttgatagtcttgcag.tg.tgccaagt.tgt.tgttgcatagag.tg.tgca.tac.tggttata.ctcta | 1978 |
| Db | 1864 | gtttcttgatagtcttgcag.tg.tgccaagt.tgt.tgttgcatagag.tg.tgca.tac.tggttata.ctcta | 1923 |

[illegible]

```

|||||
Db 2920 accatcttgatgagaccaaaagctgttacttcaactcguaaatatggaagaacatcgagg 2979
QY 3119 gctccctgtcaatctcagatggaagatccacttttaagaagaattatcgcttccatgca 3178
Db 2980 gctccctgtcaatctcagatggaagatccacttttaagaagaattatcgcttccatgca 3039
QY 3179 atcaatggtcacataatgtatgatactaccgtgttagtaatgtgtcgaagatcaagaagatt 3238
Db 3040 atcaatggtcacataatgtatgatactaccgtgttagtaatgtgtcgaagatcaagaagatt 3099
QY 3239 cgaatgtatctgtctcagcagctggtgagcaaatgtaaaacatccatctatctcaatctcagtgga 3298
Db 3100 cgaatgtatctgtctcagcagctggtgagcaaatgtaaaacatccatctatctcaatctcagtgga 3159
QY 3299 catgtgttcaactgtacgaaaaaaagagagatataaaatgtgcaactgtacaaactctctatcca 3358
Db 3160 catgtgttcaactgtacgaaaaaaagagagatataaaatgtgcaactgtacaaactctctatcca 3219
QY 3359 ggtgtttttgagacagctggaatgttaccatccaaagctgtgaatttggcggtggaatgac 3418
Db 3220 ggtgtttttgagacagctggaatgttaccatccaaagctgtgaatttggcggtggaatgac 3279
QY 3419 ctatttggcgaagatctacatgtctggatgagacacacttttctgtgtgtacagcaataag 3478
Db 3280 ctatttggcgaagatctacatgtctggatgagacacacttttctgtgtgtacagcaataag 3339
QY 3479 tgtcagaactccctggtgaatggtctctggaacacattagagatttcaagattacagcttca 3538
Db 3340 tgtcagaactccctggtgaatggtctctggaacacattagagatttcaagattacagcttca 3399
QY 3539 ggaacaatggaagaagtggtggtcccaagctgtgcagacttcatcttcccgatcaatcaat 3598
Db 3400 ggaacaatggaagaagtggtggtcccaagctgtgcagacttcatcttcccgatcaatcaat 3459
QY 3599 gctctgagacacaaaggagcccttcttctggaatcaaaagtgtgcatctgttggaccaatgatt 3658
Db 3460 gctctgagacacaaaggagcccttcttctggaatcaaaagtgtgcatctgttggaccaatgatt 3519
QY 3659 attcagcgacatcaagaacccaagggtgcccgatcagaagttctccagccttcatctctcag 3718
Db 3520 attcagcgacatcaagaacccaagggtgcccgatcagaagttctccagccttcatctctcag 3579
QY 3719 ttatcatcatgatatgtcttctgtggaagaatgtgcagacttatcgagaattccact 3778
Db 3580 ttatcatcatgatatgtcttctgtggaagaatgtgcagacttatcgagaattccact 3639
QY 3779 ggaaccttaatggtcttcttctgtgcaatgtggaatctctggaataaaacaatatctttt 3838
Db 3640 ggaaccttaatggtcttcttctgtgcaatgtggaatctctggaataaaacaatatctttt 3699
QY 3839 aaacctcaaatatgtctcgaatacatccgtttgacccaacatcatatagcattcgacg 3898
Db 3700 aaacctcaaatatgtctcgaatacatccgtttgacccaacatcatatagcattcgacg 3759
QY 3899 acctctcgataggaatgttgggtgtgttgaatgaattgagcgcatgcatgggaatg 3958
Db 3760 acctctcgataggaatgttgggtgtgttgaatgaattgagcgcatgcatgggaatg 3819
QY 3959 ggaagtaagaagaataatcgaatgacagatctactgtcttcaacttcaacaatagttt 4018
Db 3820 ggaagtaagaagaataatcgaatgacagatctactgtcttcaacttcaacaatagttt 3879
QY 4019 gccacctgtgtcttctcaaaagctcgacttccctccaagggaaggaatgactgtgagga 4078
Db 3880 gccacctgtgtcttctcaaaagctcgacttccctccaagggaaggaatgactgtgagga 3939
QY 4079 cctcaggttaataatacaaaagtgctgcaagtgtgacttccaagaagaacaatgaagtc 4138
Db 3940 cctcaggttaataatacaaaagtgctgcaagtgtgacttccaagaagaacaatgaagtc 3999
QY 4139 acagaggtactactactcaggaaggaataatctctgttaccagcatgtatgtgaagggttc 4198
|||||

```

```

Db 4000 acagagtaactactcagggatlaaaatctctgtcttaccagcaatgtatgtgaaggagtc 4059
QY 4199 ctatcttccagcagctcaagtggtccatcatgtgactctcttttccaagaatggaagta 4258
Db 4060 ctatcttccagcagctcaagtggtccatcatgtgactctcttttccaagaatggaagta 4119
QY 4259 aaggttttccaggaatacagaactccttcaacacctgtgtgacactctcagaaccccg 4318
Db 4120 aaggttttccaggaatacagaactccttcaacacctgtgtgacactctcagaaccccg 4179
QY 4319 ttactgactcgtactactcgaattcaccaccccaagatgtgggtgcaccagatgtccctgag 4378
Db 4180 ttactgactcgtactactcgaattcaccaccccaagatgtgggtgcaccagatgtccctgag 4239
QY 4379 atggaagttctgggtcgtgcagagcaggaactctac 4414
Db 4240 atggaagttctgggtcgtgcagagcaggaactctac 4275

```

## RESULT 13

AA080447  
ID AA080447 standard; DNA; 4272 BP.

XX AA080447;  
XX

DT 10-OCT-1990 (first entry)

XX DE Modified factor VIII:C sequence with the R740-D1658 deletion.

XX KW Modified factor VIII:C; haemophilia; procoagulant;

KW blood coagulation; RD deletion; ss.

OS Homo sapiens.

PN W08800831-A.

XX PD 11-FEB-1988.

XX PF 31-JUL-1987; 87WO-US01814.

XX PR 01-AUG-1986; 86US-0893375.

XX PA (BIO) BIOGEN NV.

XX PI Pask MP;

DR WPI; 1988-049866/07.

XX P-PSDB; AAP80268.

PT New DNA sequences encoding modified factor VIII:C - with deletion of DNA

XX encoding maturation polypeptide, useful for high yield transformation.

XX Claim 3; Page 47-48-49-50; 97pp; English.

CC The RD deletion removes the DNA from Ser 741 to Ser 1657. A major part

CC of the sequence encoding the maturation polypeptide of

CC factor VIII:C is deleted, i.e. Gln 744 - Asp 1563.

CC The full length factor VIII:C cDNA has two changes with respect to the

CC published sequence (EPO application 160457):

CC CTG to CTA at Leu 242 and TTC to CTC change at amino acid residue 1880

CC (Phe to Leu). The product is produced in approx. 20 times higher

CC yields than previous recombinant produced factor VIII:C and are more

CC easily purified. The peptide is used for treating haemophilia A, both

CC acute and prolonged bleeding.

XX See also AA080444 and AA080446.

SO Sequence 4272 BP; 1243 A; 941 C; 944 G; 1144 T; 0 other;

Query Match 94.4%; Score 4170; DB 9; Length 4272;  
Best Local Similarity 98.0%; Pred. No. 0;  
Matches 4267; Conservative 0; Mismatches 5; Indels 84; Gaps 1;



|    |      |   |      |
|----|------|---|------|
| QY | 59   | gccaacgaagaatactactacccctggtgtgcagttggaactgtccaatgagcatatagtcaaaagtat | 118  |
| Db | 1    | gccaacgaagaatactactacccctggtgtgcagttggaactgtccaatgagcatatagtcaaaagtat | 60   |
| QY | 119  | ctcggggaagtgcctgtggtggaacgaagaattctctctctagagtgagccaataattcttccatc    | 178  |
| Db | 61   | ctcggggaagtgcctgtggtggaacgaagaattctctctctagagtgagccaataattcttccatc    | 120  |
| QY | 179  | aacaactcagtcgtgtgtaacaanaaagacactgtgttgtaagaatltcaacggttccattccac     | 238  |
| Db | 121  | aacaactcagtcgtgtgtaacaanaaagacactgtgttgtaagaatltcaacggttccattccac     | 180  |
| QY | 239  | atcgcctaaagccaagagccaccccttggaatgtgtctgtcgtatgctccaccatccaggtcgtatg   | 298  |
| Db | 181  | atcgcctaaagccaagagccaccccttggaatgtgtctgtcgtatgctccaccatccaggtcgtatg   | 240  |
| QY | 299  | tatgtacacgttgttcatltacacttaagaataatgcttcccatctctgtacgtcttcatgct       | 358  |
| Db | 241  | tatgtacacgttgttcatltacacttaagaataatgcttcccatctctgtacgtcttcatgct       | 300  |
| QY | 359  | gtctgtgtatctccactctggaagaactctctgagagagcttgaaatatagtatgtacagaccatcaa  | 418  |
| Db | 301  | gtctgtgtatctccactctggaagaactctctgagagagcttgaaatatagtatgtacagaccatcaa  | 360  |
| QY | 419  | agggagaaagaagaatgtataaagtctctccctgtgtggaagccaatataatgtctgtgcaggtc     | 478  |
| Db | 361  | agggagaaagaagaatgtataaagtctctccctgtgtggaagccaatataatgtctgtgcaggtc     | 420  |
| QY | 479  | cgtgaagaagaatgtgtccaatgtgctctgcagcccaatgtgaccttaactactcaactcttct      | 538  |
| Db | 421  | cgtgaagaagaatgtgtccaatgtgctctgcagcccaatgtgaccttaactactcaactcttct      | 480  |
| QY | 539  | catgtgtgaacctgtgtlaaagaactctgaattctcagccctcatcttggaagccctactagtatgata | 598  |
| Db | 481  | catgtgtgaacctgtgtlaaagaactctgaattctcagccctcatcttggaagccctactagtatgata | 540  |
| QY | 599  | gaagggagttctgtgcgaaggaagaacagacagacctgtgcacaataattatactacttttgtct     | 658  |
| Db | 541  | gaagggagttctgtgcgaaggaagaacagacagacctgtgtgcacaataattatactacttttgtct   | 600  |
| QY | 659  | gaatttggatggaaggggaanaagtgtgcacatcgaagaacaaagaactcttggatgtcagaataga   | 718  |
| Db | 601  | gaatttggatggaaggggaanaagtgtgcacatcgaagaacaaagaactcttggatgtcagaataga   | 660  |
| QY | 719  | gaatgtcgaactctgtctcggggctgtgccttaaaatgtgcacacagtcgaatgtgtatgtlaaacaag | 778  |
| Db | 661  | gaatgtcgaactctgtctcggggctgtgccttaaaatgtgcacacagtcgaatgtgtatgtlaaacaag | 720  |
| QY | 779  | tctctgtccaggtctctgtatctgtaatgtgcacacagaagaatcagttcatgtgcatgtgtggaatg  | 838  |
| Db | 721  | tctctgtccaggtctctgtatctgtaatgtgcacacagaagaatcagttcatgtgcatgtgtggaatg  | 780  |
| QY | 839  | ggcacaacaccctgaagaatgtgcacaaataatctctcgaagaatgcacaacattctctgtgagaac   | 898  |
| Db | 781  | ggcacaacaccctgaagaatgtgcacaaataatctctcgaagaatgcacaacattctctgtgagaac   | 840  |
| QY | 899  | catctgcgaagcgtctccttggnaatctcgcgcacaataacttctctactgtgtccaacactcttg    | 958  |
| Db | 841  | catctgcgaagcgtctccttggnaatctcgcgcacaataacttctctactgtgtccaacactcttg    | 900  |
| QY | 959  | atggaaccttggaagacttctctactcgtttgtgtcatatactctctccacaacaaatgtatgcatg   | 1018 |
| Db | 901  | atggaaccttggaagacttctctactcgtttgtgtcatatactctctccacaacaaatgtatgcatg   | 960  |
| QY | 1019 | gaagacttatgtcaaaatgtacagactgtgcgaaggaacccaacatcagaatggaanaataat       | 1078 |
| Db | 961  | gaagacttatgtcaaaatgtacagactgtgcgaaggaacccaacatcagaatggaanaataat       | 1020 |
| QY | 1079 | gaagaagcggaaagaactatgtatgtatccttacttaccgtatcttgaatgtgtgtgtcgaagttt    | 1138 |
| Db | 1021 | gaagaagcggaaagaactatgtatgtatccttacttaccgtatcttgaatgtgtgtgtcgaagttt    | 1080 |
| QY | 1139 | gatgatgacaactctcccttccattatccaataattcgtctcagttgtgcagaagaacatctcaaa    | 1198 |

|    |      |   |      |
|----|------|---|------|
| Db | 1081 | gatgataaacctccctcccttcccttaacaaatctgcagcttgccaaagaacatccctaa      | 1140 |
| Oy | 1199 | acttggacaattcaatgctgcgtcgaaagagagactcgggacatgctcccttaagcttc       | 1258 |
| Db | 1141 | acttggtaacttcaatctgctgcgtcgaaagagagactcgggacatgctcccttaagcttc     | 1200 |
| Oy | 1259 | gccccgaatgacagaagttataaaagtcaattctgaacaatgcccctcagcgatctgt        | 1318 |
| Db | 1201 | gccccgaatgacagaagttataaaagttcaattctgaacaatgcccctcagcgatctgt       | 1260 |
| Oy | 1319 | aggaaatcaaaaaaagtcgcattatgctataccgataagaaaccttaagctcgttaa         | 1378 |
| Db | 1261 | aggaaatcaaaaaaagtcgcattatgctataccgataagaaaccttaagctcgttaa         | 1320 |
| Oy | 1379 | gctatccagcatgatacgaagatctggagcccttcaattatgaggaaagtctgagacaca      | 1438 |
| Db | 1321 | gctatccagcatgatacgaagatctggagcccttcaattatgaggaaagtctgagacaca      | 1380 |
| Oy | 1439 | ctgttgattatatttaagaatcaagaagacacataacatctacccctcaggaatc           | 1498 |
| Db | 1381 | ctgttgattatatttaagaatcaagaagacacataacatctacccctcaggaatc           | 1440 |
| Oy | 1499 | actgataccgctctcttgattatcaagaagttaccaaaggtgtataaacatttgaagat       | 1558 |
| Db | 1441 | actgataccgctctcttgattatcaagaagttaccaaaggtgtataaacatttgaagat       | 1500 |
| Oy | 1559 | ttccaattctgcaggaagaataattcaaatataatggacagctgactgtagaagaatgg       | 1618 |
| Db | 1501 | ttccaattctgcaggaagaataattcaaatataatggacagctgactgtagaagaatgg       | 1560 |
| Oy | 1619 | ccaactaaatcagatctctcggtgcgtgacccgctattactctagtcttgtaataagag       | 1678 |
| Db | 1561 | ccaactaaatcagatctctcggtgcgtgacccgctattactctagtcttgtaataagag       | 1620 |
| Oy | 1679 | aggagatcagcttcaagaatcatttgagccctctccatctgctatacaagaagatctgtaat    | 1738 |
| Db | 1621 | aggagatcagcttcaagaatcatttgagccctctccatctgctatacaagaagatctgtaat    | 1680 |
| Oy | 1739 | caaaaggaagaaccgaataatgtcagacaagaagaatgctacccgtttctctgattgat       | 1798 |
| Db | 1681 | caaaaggaagaaccgaataatgtcagacaagaagaatgctacccgtttctctgattgat       | 1740 |
| Oy | 1799 | gagaacccgaagctgtgtacctccacagagaatatacaacgctctctcccaatccagctgga    | 1858 |
| Db | 1741 | gagaacccgaagctgtgtacctccacagagaatatacaacgctctctcccaatccagctgga    | 1800 |
| Oy | 1859 | gtgcagctcttgaggttccaaggtttccaagctccaaatcatgtcagacagcataaatgtgcat  | 1918 |
| Db | 1801 | gtgcagctcttgaggttccaaggtttccaagctccaaatcatgtcagacagcataaatgtgcat  | 1860 |
| Oy | 1919 | gtttttgtaagtcttgcaagttgtcagttgtcttgcatgtgaggtgcatatcgtgatactcta   | 1978 |
| Db | 1861 | gtttttgtaagtcttgcaaggtttgtcagttgtcttgcatgtgaggtgcatatcgtgatactcta | 1920 |
| Oy | 1979 | agcatttggacagacagactgacattcccttctgtcctctctctctgtgatactcctcaaac    | 2038 |
| Db | 1921 | agcatttggacagacagactgacattcccttctgtcctctctctctgtgatactcctcaaac    | 1980 |
| Oy | 2039 | aaaatgtgctatgtgaagacacactccacctattcccatctcaaggagaaactgtctcatg     | 2098 |
| Db | 1981 | aaaatgtgctatgtgaagacacactccacctattcccatctcaaggagaaactgtctcatg     | 2040 |
| Oy | 2099 | tcgatatgaaaaccccaagttcatatgattctcgtgggtgcacaacatccaaacttcgggaacga | 2158 |
| Db | 2041 | tcgatatgaaaaccccaagttcatatgattctcgtgggtgcacaacatccaaacttcgggaacga | 2100 |
| Oy | 2159 | ggcatgaccgctcttaccgaaggtttctagtcttgacaaagaacactgtgtattattacag     | 2218 |
| Db | 2101 | ggcatgaccgctcttaccgaaggtttctagtcttgacaaagaacactgtgtattattacag     | 2160 |
| Oy | 2219 | gacagttatgaagatattccaagcactgtgctgataaaacaatgccaatttgaaccaaga      | 2278 |



Dh 2161 gacagatataagataattcagacactctgctagtaaaacaatgcatatgaaccaaga 2220  
Qy 2279 agcttccagagattccaagaacacctagcaactagcaaaagaatttaatgcaccoca 2338  
Db 2221 ----- 2220  
Qy 2339 ccagcttgaaacgcacatcaaacggaataaactcgtactactcttcagtcaga tcaag 2398  
2221 ----- gaataaactcgtactactcttcagtcaga tcaag 2256  
Qy 2399 gaattgactatgtatataccatacaagttgaatagaagaagaatttgcattat 2458  
2257 gaattgactatgtatataccatacaagttgaatagaagaagaatttgcattat 2316  
Db 2459 gatgagatgaataatcagaagcccgccagcttcacaagaacaacagacactttat 2518  
2317 gatgagatgaataatcagaagcccgccagcttcacaagaacaacagacactttat 2376  
Qy 2519 gctcagatgagagagctcctggaatataagatagtagtccccacatgcttaagaac 2578  
2377 gctcagatgagagagctcctggaatataagatagtagtccccacatgcttaagaac 2436  
Db 2579 aaggctcagatgagagctcctcagctcacaagaagaagttgtttccagaatttactat 2638  
2437 aaggctcagatgagagctcctcagctcacaagaagaagttgtttccagaatttactat 2496  
Qy 2639 ggcctcttactcagcccttatccgttgagaactaaatgaagaatttggactccggg 2698  
2497 ggcctcttactcagcccttatccgttgagaactaaatgaagaatttggactccggg 2556  
Db 2699 ccataataagagcaagaattgaagaataatcatatgtaacttccagaatcagccct 2758  
2557 ccataataagagcaagaattgaagaataatcatatgtaacttccagaatcagccct 2616  
Qy 2759 cgtccctatctctatctatagccttattcttaagagaagatcagaagcagaagca 2818  
2617 cgtccctatctctatctatagccttattcttaagagaagatcagaagcagaagca 2676  
Db 2819 gaaccctagaataaacttctgaagccttaatgaacccaataacttcttggaaagcgaa 2878  
2677 gaaccctagaataaacttctgaagccttaatgaacccaataacttcttggaaagcgaa 2736  
Qy 2879 catcataatgcaaccctaaagaatgagtttgactgcaaaagcctggtctatctctga 2938  
2737 catcataatgcaaccctaaagaatgagtttgactgcaaaagcctggtctatctctga 2796  
Db 2939 gttgacctggaataaagaatgtaacatcagagcctgattggaacccctctgctgcaact 2998  
2797 gttgacctggaataaagaatgtaacatcagagcctgattggaacccctctgctgcaact 2856  
Qy 2999 aacacactgaacccctgctatgttgagagaagaagtagacagtagaagaatttgcctgttttc 3058  
2857 aacacactgaacccctgctatgttgagagaagaagtagacagtagaagaatttgcctgttttc 2916  
Qy 3059 accatcttggatgagacccaagaagctgttacttcaactgaataatataagaaagaaactgcag 3118  
2917 accatcttggatgagacccaagaagctgttacttcaactgaataatataagaaagaaactgcag 2976  
Db 3119 gctccctgcaataatccagaatgaagaatcccaactttaaagaataatcatcgtctcactga 3178  
2977 gctccctgcaataatccagaatgaagaatcccaactttaaagaataatcatcgtctcactga 3036  
Qy 3179 atcaaatggtatataatgtagatcacactctgtttagttagtgcctagaatcaagaat 3238  
3037 atcaaatggtatataatgtagatcacactctgtttagttagtgcctagaatcaagaat 3096  
Qy 3239 cgaatgtagtctcagcagctgagcaagaatgaagaacatcatctctatcaatcagtaga 3298  
3097 cgaatgtagtctcagcagctgagcaagaatgaagaacatcatctctatcaatcagtaga 3156  
Db 3299 catgttctactgtagcagaaaaaagagagtagataaagtggaactgtacaatctctatca 3358  
3157 catgttctactgtagcagaaaaaagagagtagataaagtggaactgtacaatctctatca 3216

Qy 3359 ggtgttttgagacagtggaatggtttaccatcccaagctggaaatttggcggtggaatgc 3418  
3217 ggtgttttgagacagtggaatggtttaccatcccaagctggaaatttggcggtggaatgc 3276  
Db 3419 ctatattggcagatctacatgcttgagatgagcacacttcttctggtgtacagcaataag 3478  
3277 ctatattggcagatctacatgcttgagatgagcacacttcttctggtgtacagcaataag 3336  
Qy 3479 tgtcagaactcccttggaatggttcttgacacaaatagaagaatttcaagatcaagcttca 3538  
3337 tgtcagaactcccttggaatggttcttgacacaaatagaagaatttcaagatcaagcttca 3396  
Db 3539 ggaacaatggaagatgggcccccaagcttgagcagactcatatccgatcaatcaat 3598  
3397 ggaacaatggaagatgggcccccaagcttgagcagactcatatccgatcaatcaat 3456  
Qy 3599 gcttgagacccaagaagccttcttcttgatcaaggttgatctgtttggcaccatgatt 3658  
3457 gcttgagacccaagaagccttcttcttgatcaaggttgatctgtttggcaccatgatt 3516  
Db 3659 attcagcgaatcaagaaccaaggttgcctgacagaagttctcagcctctacatctcag 3718  
3517 attcagcgaatcaagaaccaaggttgcctgacagaagttctcagcctctacatctcag 3576  
Qy 3719 ttatcatatgtagtagtcttgaatggagaagtggaagattatcagaagaatttccact 3778  
3577 ttatcatatgtagtagtcttgaatggagaagtggaagattatcagaagaatttccact 3636  
Db 3779 ggaaccctaaatgcttctcttgacaaatgtagatctcatcgggataaacaataatctt 3838  
3637 ggaaccctaaatgcttctcttgacaaatgtagatctcatcgggataaacaataatctt 3696  
Qy 3839 aacctccaatattatgctcgaatataatccgtttgacccaactcaatataagatctgcag 3898  
3697 aacctccaatattatgctcgaatataatccgtttgacccaactcaatataagatctgcag 3756  
Db 3899 actctcgaatgagatgtagtggcgttgatataaagttagttgacagatctgcagatggaa 3958  
3757 actctcgaatgagatgtagtggcgttgatataaagttagttgacagatctgcagatggaa 3816  
Qy 3959 gagagtaagaacaaatcaatagatgacagatattactgcttcaatcattcaatcaatgatt 4018  
3817 gagagtaagaacaaatcaatagatgacagatattactgcttcaatcattcaatcaatgatt 3876  
Db 4019 gccacctggtctcttccaagaagctgacttcaactccaagggagagtagtccctgaga 4078  
3877 gccacctggtctcttccaagaagctgacttcaactccaagggagagtagtccctgaga 3936  
Qy 4079 cctcaggtgaataatccaaaaggtgctgcaagtggaacttccaagaagaatgaagaatgc 4138  
3937 cctcaggtgaataatccaaaaggtgctgcaagtggaacttccaagaagaatgaagaatgc 3996  
Qy 4139 acagagtagtaactactcagaggttaaatctctgcttaccagcatgtagtgaagaatgc 4198  
3997 acagagtagtaactactcagaggttaaatctctgcttaccagcatgtagtgaagaatgc 4056  
Db 4199 ctcaatctcagcagatcaatgagcctcagctggaactcttcttccaagaatggaagaatga 4258  
4057 ctcaatctcagcagatcaatgagcctcagctggaactcttcttccaagaatggaagaatga 4116  
Qy 4259 aaggttttccaaggaataaagaactctcacaacgttggaactcttaagaaccaagc 4318  
4117 aaggttttccaaggaataaagaactctcacaacgttggaactcttaagaaccaagc 4176  
Db 4319 ttactgactcgttacccttgaatcaaccccaagagtttggtgacacagaattgcctcagag 4378  
4177 ttactgactcgttacccttgaatcaaccccaagagtttggtgacacagaattgcctcagag 4236  
Qy 4379 atggaagttcttggtctgagagcaagaactctac 4414  
4237 atggaagttcttggtctgagagcaagaactctac 4272  
Db

RESULT 14  
AA82258  
ID AAX82258 standard; cDNA; 4373 BP.  
XX  
AC AAX82258;  
XX  
DT 18-AUG-1999 (first entry)  
XX  
DE Beta domain deleted Factor VIII protein encoding cDNA.  
XX  
KW Factor VIII protein; gene modification; gene therapy; clinical disorder;  
KW splicing pattern; RNA processing; gene regulation; beta-domain; human;  
KW ss.  
XX  
OS Homo sapiens.  
XX  
PN W09929848-A1.  
XX  
PD 17-JUN-1999.  
XX  
PF 25-NOV-1998; 98MO-US25354.  
XX  
PR 16-JAN-1998; 98US-0071596.  
PR 05-DEC-1997; 97US-0067614.  
XX  
PA (IMMU-) IMMUNE RESPONSE CORP.  
XX  
PI Bidlingmaier S, Gonzales JEN, Ill CR, Yang CQ;  
XX  
DR WPI; 1999-385602/32.  
DR P-PSDB; AAY21675.  
XX  
PT Genes and vectors exhibiting increased expression and novel splicing  
PT patterns, useful for expression of, e.g. beta-domain deleted factor  
PT VIII  
XX  
PS Claim 9; Page 72-78; 123pp; English.  
XX  
PS The invention describes novel genes and vectors exhibiting increased  
XX expression and novel splicing patterns. It provides a gene encoding a  
XX Factor VIII protein, that comprises one or more consensus or near  
XX consensus splice sites which have been corrected to increase expression.  
XX The method, DNA sequences and expression vectors can be used to increase  
XX the expression of a gene, especially a Factor VIII gene. Genes containing  
XX modified 5' and/or 3' untranslated regions have optimized expression  
XX levels and tissue-specific expression. The methods are used for  
XX identification and correction of consensus splice sites, addition of  
XX introns, optimization of 5' and 3' untranslated regions and increase in  
XX cytoplasmic RNA accumulation. Hence the DNAs are useful in gene therapy  
XX to treat a clinical disorder, to study RNA processing and/or gene  
XX regulation. The present sequence represents a cDNA encoding a beta-domain  
XX deleted factor VIII protein.  
SQ Sequence 4373 BP; 1278 A; 1011 C; 938 G; 1146 T; 0 other:  
  
Query Match 94.2%; Score 4161; DB 20; Length 4373;  
Best Local Similarity 96.8%; Pred. No. 0;  
Matches 42/3; Conservative 0; Mismatches 100; Indels 42; Gaps 1;

QY 182 acctcagtcgtgtacaaaagactctgtgtgtagaattcaagcttcacacctttcaacatc 241  
DB |||||||  
QY 181 acctcagtcgtgtacaaaagactctgtgtgtagaattcaagcttcacacctttcaacatc 240  
DB |||||||  
QY 242 gctaaagcgaagccacctgtgagtgtctgtctaggtctccatccagctcaggtttat 301  
DB |||||||  
QY 241 gctaaagcgaagccacctgtgagtgtctgtctaggtctccatccagctcaggtttat 300  
DB |||||||  
QY 302 gatacagtgatctataccttaagaacatgctccctccatccctgtcagcttcctgtgt 361  
DB |||||||  
QY 301 gatacagtgatctataccttaagaacatgctccctccatccctgtcagcttcctgtgt 360  
DB |||||||  
QY 362 ggtgtatcctactgtgaaagctctctgaggagctgtgaatatgatacagacagtcacaag 421  
DB |||||||  
QY 361 ggtgtatcctactgtgaaagctctctgaggagctgtgaatatgatacagacagtcacaag 420  
DB |||||||  
QY 422 gagaagaagaatgataaagctctccctgtgtggaagccatacatatgtctgcaagctcgt 481  
DB |||||||  
QY 421 gagaagaagaatgataaagctctccctgtgtggaagccatacatatgtctgcaagctcgt 480  
DB |||||||  
QY 482 aaagaagaatgctccaatgtgaccttgaccacctgtgtccttaacctatactttctcat 541  
DB |||||||  
QY 481 aaagaagaatgctccaatgtgaccttgaccacctgtgtccttaacctatactttctcat 540  
DB |||||||  
QY 542 gtgacactgtgtaaaaagacttgaaattcagagcctcatgtgagcctactagatgtagaaga 601  
DB |||||||  
QY 541 gtgacactgtgtaaaaagacttgaaattcagagcctcatgtgagcctactagatgtagaaga 600  
DB |||||||  
QY 602 gggagctcgtgccaagaaagaaacacagaccttggacaatatattactactcttctgtcta 661  
DB |||||||  
QY 601 gggagctcgtgccaagaaagaaacacagaccttggacaatatattactactcttctgtcta 660  
DB |||||||  
QY 662 ttgatgaaggaaagtgtgacatccagaacaaagacctcctgtatgcagataagat 721  
DB |||||||  
QY 661 ttgatgaaggaaagtgtgacatccagaacaaagacctcctcatcagacagataagat 720  
DB |||||||  
QY 722 gctgcatctgctcggtcggtggtcctaataatgacacagtcacatgttatgtataacaggtct 781  
DB |||||||  
QY 721 gctgcatctgctcggtcggtggtcctaataatgacacagtcacatgttatgtataacagga 780  
DB |||||||  
QY 782 ctgccaagctcgtatgtatgacagagaatcagtcattgtgattgattgtatgattgagtc 841  
DB |||||||  
QY 781 ctgccaagctcgtatgtatgacagagaatcagtcattgtgattgattgtatgattgagtc 840  
DB |||||||  
QY 842 acctcctctgaaagcactcaatatctcctcgaagtcacacatctctgtgaaagacact 901  
DB |||||||  
QY 841 acctcctctgaaagcactcaatatctcctcgaagtcacacatctctgtgttgaataacct 900  
DB |||||||  
QY 902 cgcgaagcgtctctgtgaatctgcgaataacttctcctactcgtcgaacactctgtatg 961  
DB |||||||  
QY 901 cgcgaagcgtctctgtgaatctgcgaataacttctcctactcgtcgaacactctctcatg 960  
DB |||||||  
QY 962 gaccttggacagttctctactcgttttgcataatctctccacacaatgattgcaaggaa 1021  
DB |||||||  
QY 961 gaccttggacagttctctactcgttttgcataatctctccacacaatgattgcaaggaa 1020  
DB |||||||  
QY 1022 gcttatgtcaaaagttagacaagctgtccagaggaaccccaactacgaatgaaaataatgaa 1081  
DB |||||||  
QY 1021 gcttatgtcaaaagttagacaagctgtccagaggaaccccaactacgaatgaaaataatgaa 1080  
DB |||||||  
QY 1082 gaagcgaagactatgatagtatctactgattctgaaatgaaatgagtgtcaggtttgat 1141  
DB |||||||  
QY 1081 gaagcgaagactatgatagtatctactgattctgaaatgaaatgagtgtcaggtttgat 1140  
DB |||||||  
QY 1142 gatgaacactctcctctcttatccaatcgcgtcagttgccaagaagactcctaanaact 1201  
DB |||||||  
QY 1141 gatgaacactctcctctcttatccaatcgcgtcagttgccaagaagactcctaanaact 1200  
DB |||||||  
QY 1202 ttggtacattacattctgtctgaaagaggaagctggaacatctcccttaagctccgcgc 1261  
DB |||||||  
QY 1201 ttggtacattacattctgtctgaaagaggaagctggaacatctcccttaagctccgcgc 1260  
DB |||||||  
QY 1262 ccgcatgacagaagtataaaagtcaatatattgaaacaatggtccctcagcggattgtgtag 1321

|||||  
Db 1261 ccgagtacagaagttataaagtcataatttgacacatgcccctcagcgattggaag 1320  
QY 1322 aagtacaaaaagtcgatttatgtgcatacagatgaaccccttgaagctctgagct 1381  
Db 1321 aagttacaaaaagtcgatttatgtgcatacagatgaaccccttgaagctctgagct 1380  
QY 1382 attcagatgaatcagaagaaacttggagcccttcttattatggggaagtggagcacatg 1441  
Db 1381 attcagatgaatcagaagaaacttggagcccttcttattatggggaagtggagcacatg 1440  
QY 1442 ttgattatattaagaatacagaagcagacacataacatactacccctcagcaatcacc 1501  
Db 1441 ctcatattatattaagaatacagaagcagacacataacatactacccctcagcaatcacc 1500  
QY 1502 gatgtccgtcctctgtattcttaagaagattcccaaaagtgtaaaactttgaagatctt 1561  
Db 1501 gatgtccgtcctctgtattcttaagaagattcccaaaagtgtaaaactttgaagatctt 1560  
QY 1562 ccaattctgcagagagaatcattcaaatataaattgacagtgactgtagaagaattggcca 1621  
Db 1561 ccaattctgcagagagaatcattcaaatataaattgacagtgactgtagaagaattggcca 1620  
QY 1622 actaaatcagatccctcgtgtgcctgcagccgctatctactagtcttgaatagagaga 1681  
Db 1621 actaaatcagatccctcgtgtgcctgcagccgctatctactagtcttgaatagagaga 1680  
QY 1682 gatctagtctcagagctcatctatgtgcctctctctatctgtcacaagaagatctgagatcaa 1741  
Db 1681 gatctagtctcagagctcatctatgtgcctctctctatctgtcacaagaagatctgagatcaa 1740  
QY 1742 agagggaacacagataatgtcagacagaagaaatgcatacctgtcttctgtattgtatgag 1801  
Db 1741 agagggaacacagataatgtcagacagaagaaatgcatacctgtcttctgtattgtatgag 1800  
QY 1802 aaccggaagctgtgtactcctcacagaaatatacaacgcttctccccaatccagcttggagt 1861  
Db 1801 aaccggaagctgtgtactcctcacagaaatatacaacgcttctccccaatccagcttggagt 1860  
QY 1862 cagcttgagagatccagatgtccaaagctccaaacatactgacagcatcaatgtgtatgtc 1921  
Db 1861 cagcttgagagatccagatgtccaaagctccaaacatactgacagcatcaatgtgtatgtc 1920  
QY 1922 ttgtatagtttgagctgtcagcttgtgttcgactgagtggtgacacgtgacatcttaagc 1981  
Db 1921 ttgtatagtttgagctgtcagcttgtgttcgactgagtggtgacacgtgacatcttaagc 1980  
QY 1982 attgagacagagactgactcct 2041  
Db 1981 attgagacagagactgactcct 2040  
QY 2042 atgtgtctatagaacacacacaccctcattcccaattctcagagagaacatgtctctcatgtcg 2101  
Db 2041 atgtgtctatagaacacacacaccctcattcccaattctcagagagaacatgtctctcatgtcg 2100  
QY 2102 atggaacacacagctctatgactctggtgtgcacaaactcagactctggaacagagc 2161  
Db 2101 atggaacacacagctctatgactctggtgtgcacaaactcagactctggaacagagc 2160  
QY 2162 atgacgctctactgaaggttctctagttgtgacaagaacacatggtatattaacgagac 2221  
Db 2161 atgacgctctactgaaggttctctagttgtgacaagaacacatggtatattaacgagac 2220  
QY 2222 agttatgaagatatcttcagatatctgtctgtagtaaaaaacatgccaattgacaagaagc 2281  
Db 2221 agttatgaagatatcttcagatatctgtctgtagtaaaaaacatgccaattgacaagaagc 2280  
QY 2282 ttctccagaaattcaagacacccctagacatgcaaaaagaatttaatgacccccca 2341  
Db 2281 ttctccagaaattcaagacacccctagacatgcaaaaagaatttaatgacccccca 2340  
QY 2342 gtcttgaacgcatacaacgggaataaactcgtactactcttcagtcagatcaagaagaa 2401  
Db 2341 gtcttgaacgcatacaacgggaataaactcgtactactcttcagtcagatcaagaagaa 2400

Db 2299 gtcttgaacgcatacaacgggaataaactcgtactactcttcagtcagatcaagaagaa 2358  
QY 2402 attgactatgatataccatatcagtttgaaatgaagaagaatttgcatttatgat 2461  
Db 2359 attgactatgatataccatatcagtttgaaatgaagaagaatttgcatttatgat 2418  
QY 2462 gagatgaanaatcagaagcccccgacgttccaaagaacacagacatttctatgtc 2521  
Db 2419 gagatgaanaatcagaagcccccgacgttccaaagaacacagacatttctatgtc 2478  
QY 2522 gcaatgagagagctctgagattatggagatgagctccccaatgcttctaagaagaag 2581  
Db 2479 gcaatgagagagctctgagattatggagatgagctccccaatgcttctaagaagaag 2538  
QY 2582 gctcagagatgagctctcctcagttccaagaagttgtttccaagaatttactgagc 2641  
Db 2539 gctcagagatgagctctcctcagttccaagaagttgtttccaagaatttactgagc 2598  
QY 2642 tcccttactcagccctatataccgttgagaaactaaatgaacatttggagctcctggagcca 2701  
Db 2599 tcccttactcagccctatataccgttgagaaactaaatgaacatttggagctcctggagcca 2658  
QY 2702 tatataagacagaagttgagataataatcaatggttaacttcagaaatcagccctcgt 2761  
Db 2659 tatataagacagaagttgagataataatcaatggttaacttcagaaatcagccctcgt 2718  
QY 2762 cccattctccttactatctagccttattcttctatgagaagaatacagaagcagaagcaaa 2821  
Db 2719 cccattctccttactatcttctcctcctatatacatalgaggaagatcagaagcagaagcaaa 2778  
QY 2822 cctagaanaaactctgtcaagcctaaatgaacacaaacttactcttgaagaatgcaaat 2881  
Db 2779 cctagaanaaactctgtcaagcctaaatgaacacaaacttactcttgaagaatgcaaat 2838  
QY 2882 cataatgaccccaataaagaatgagcttgacgtcagaagcctggcttactctctatgtc 2941  
Db 2839 cataatgaccccaataaagaatgagcttgacgtcagaagcctggcttactctctatgtc 2898  
QY 2942 gacctggaanaaagaatgtgcactgagcctgtatgtggcccttctgtctgcacacaaac 3001  
Db 2899 gacctggaanaaagaatgtgcactgagcctgtatgtggcccttctgtctgcacacaaac 2958  
QY 3002 acactgaacccctgtcatgtgagagaagaatgacagatcaggaattctcgtctttccac 3061  
Db 2959 acactgaacccctgtcatgtgagagaagaatgacagatcaggaattctcgtctttccac 3018  
QY 3062 atcttctgataagaccaaagaatggttactcactgaatgaataatggaagaacatgcaagct 3121  
Db 3019 atcttctgataagaccaaagaatggttactcactgaatgaataatggaagaacatgcaagct 3078  
QY 3122 cccctgcaatatccagatggaagatcccaactttaaagaagaattatcgtctccatgcacac 3181  
Db 3079 cccctgcaatatccagatggaagatcccaactttaaagaagaattatcgtctccatgcacac 3138  
QY 3182 aatggtctacataatgataacacactggtctgtagtaatggtcagaagacaaagaatgcga 3241  
Db 3139 aatggtctacataatgataacacactggtctgtagtaatggtcagaagacaaagaatgcga 3198  
QY 3242 tggatctctgcacgaatgggcagcaaatgaaacatccattctatcattcagtggaacat 3301  
Db 3199 tggatctctgcacgaatgggcagcaaatgaaacatccattctatcattcctccggaacat 3358  
QY 3302 gtgttcaactgtacgaaaaaagaagaggtataaaatggcactgtgacaatcctctcaaggt 3361  
Db 3259 gtgttcaactgtacgaaaaaagaagaggtataaaatggcactgtgacaatcctctcaaggt 3318  
QY 3362 gtcttgaagacagttgaaatgcttaccataccaaagctggaatcttggcgggtgagatgcct 3421  
Db 3319 gtcttgaagacagttgaaatgcttaccataccaaagctggaatcttggcgggtgagatgcct 3378  
QY 3422 attgacgagacatcatatcgtctgagatgagcaacatttctctgtgtatcagcaataagttg 3481  
Db 3379 attgacgagacatcatatcgtctgagatgagcaacatttctctgtgtatcctcaataagttg 3438

|    |      |   |      |
|----|------|---|------|
| QY | 3482 | cagactcccttgggaatggtctcttgacacataagagatttcaagattacagcttcaagga           | 3541 |
| Db | 3439 | cgactccctctgggaatgtctcttgacacataatagaatttccagatttaccagcttccagga         | 3498 |
| QY | 3542 | caataggagcagtggggcccccaagctctgcagagattatatttccggtatcaatgccc             | 3601 |
| Db | 3499 | caataggagcagtggggcccccaagctctgcagagattatatttccggtatcaatgccc             | 3558 |
| QY | 3602 | tggagaccacaaggagcccttctcttgatcaaggctgatacgtcttgcagccaatgatatt           | 3661 |
| Db | 3559 | tggagaccacaaggagcccttctcttgatcaaggctgatacgtcttgcagccaatgatatt           | 3618 |
| QY | 3662 | caagcagcatcaagaagcccaagggtgccgttcagaaagtcttccagccctctacatctccagttt      | 3721 |
| Db | 3619 | caagcagcatcaagaagcccaagggtgccgttcagaaagtcttccagccctctacatctccaat        | 3678 |
| QY | 3722 | atcatcatgatagctcttgaatgggaagaagtgcgagacttatcgagaagaattccacgtga          | 3781 |
| Db | 3679 | atcatcatgatagctcttgaatgggaagaagtgcgagacttatcgagaagaattccacgtga          | 3738 |
| QY | 3782 | accttaatgtgctctcttcttgacaaatgttgatcacttggaataaaacacaatatttttaac         | 3841 |
| Db | 3739 | acctcatgtcttcttcttggcaaatgtgattcacttggaataaaacacaatatttttaac            | 3798 |
| QY | 3842 | ccttccaattatgtctctgatatacatccggtttgcagcccaactatataagatctgcagcaact       | 3901 |
| Db | 3799 | ccttccaattatgtctctgatatacatccggtttgcagcccaactatataagatctgcagcaact       | 3858 |
| QY | 3902 | cttcgcaatgagtttgatgggctctgatttaaatggttgcagcatcttgagaaatggag             | 3961 |
| Db | 3859 | cttcgcaatgagtttgatgggctctgatttaaatggttgcagcatcttgagaaatggag             | 3918 |
| QY | 3962 | agtaaaagcacaatacagatgcacagagattactgcttcatcctcaatttacaatatgttgc          | 4021 |
| Db | 3919 | agtaaaagcacaatacagatgcacagagattactgcttcatcctcaatttacaatatgttgc          | 3978 |
| QY | 4022 | acctgtgtctcttcaaaaagctcgactcaactcctccaaggaggaataatgctctgagacct          | 4081 |
| Db | 3979 | acctgtgtctcttcaaaaagctcgactcaactcctccaaggaggaataatgctctgagacct          | 4038 |
| QY | 4082 | caggctgaataatcccaaaagggtgcctgcaggttgagacttccagaagaacaatgaaagtacaa       | 4141 |
| Db | 4039 | caggctgaataatcccaaaagggtgcctgcaggttgagacttccagaagaacaatgaaagtacaa       | 4098 |
| QY | 4142 | ggagtaactactcaaggagagtaaaatcctctgtcttaccagacatgttaatgaaagagtttcttc      | 4201 |
| Db | 4099 | ggagtaactactcaaggagagtaaaatcctctgtcttaccagacatgttaatgaaagagtttcttc      | 4158 |
| QY | 4202 | acctcagcagatcaagaatgagcctatggtgactctcttcttttcgaattgcaagaatgaag          | 4261 |
| Db | 4159 | acctcagcagatcaagaatgagcctatggtgactctcttcttttcgaattgcaagaatgaag          | 4218 |
| QY | 4262 | gtctttcaggagaaatcaagaactccttcaacacgttggatgaactctcagaccaccgtta           | 4321 |
| Db | 4219 | gtctttcaggagaaatcaagaactccttcaacacgttggatgaactctcagaccaccgtta           | 4278 |
| QY | 4322 | ctgactcgtctacaccttcgaaattcaccccccaagatgttgggtgcacacagatctgcaccttgaagatg | 4381 |
| Db | 4279 | ctcactcgtctacaccttcgaaattcaccccccaagatgttgggtgcacacagatctgcaccttgaagatg | 4338 |
| QY | 4382 | gaggtctctgggtctgcaggtcacagaccttactg                                     | 4416 |
| Db | 4339 | gaggtctctgggtctgcaggtcacagaccttactg                                     | 4373 |

|         |                                 |
|---------|---------------------------------|
| RESULT  | 15                              |
| AA81544 |                                 |
| ID      | AA81544 standard; DNA; 4830 BP. |
| XX      |                                 |
| AC      |                                 |
| XX      | AA81544;                        |
| DT      | 04-DEC-1990 (first entry)       |

XX Human Factor VIII-C analog having exon 14 deleted.  
DE  
XX  
XX Human Factor VIII-C analog; exon deletion; coagulation disorders;  
KW haemophilia; ss.  
XX  
XX  
PN EP265778-A.  
XX  
PD 04-MAY-1988.  
XX  
XX 14-OCT-1987; 87EP-0115043.  
PF  
XX  
PR 15-OCT-1986; 86US-0919153.  
XX  
PA (RORE ) RORER INT OVERSEAS.  
XX  
PI Sarver N, Drohan W;  
XX  
DR WP1; 1988-120930/18.  
XX  
XX  
PT Human Factor VIII-C analogue free of other proteins - produced from  
XX recombinant DNA and used in the treatment of coagulation disorders.  
XX  
XX  
PS Claim 3; Page 13; 42pp; English.  
XX  
XX  
CC The genetically engineered analogue can provide a dependable and  
CC readily available therapeutic agent to be used in the treatment of  
CC haemophilia and coagulation disorders in humans or animals.  
CC See also AANB1543-45.  
CC  
XX  
SQ Sequence 4830 BP; 1392 A; 1091 C; 1058 G; 1289 T; 0 other;

|                            |       |   |       |                     |
|----------------------------|-------|---|-------|---------------------|
| Query Match                | 94.0% | Score 4152;   | DB 9; | Length 4830;        |
| Best Local Similarity      | 95.0% | Pred. No. 0;  |       |                     |
| Matches 4381; Conservative | 0;    | Mismatches  | 35;   | Indels 198; Gaps 1. |
| QY                         | 2     | atgcaataaagctctccacctgctctcttcctctgctgctcttgcgattcgtcttaagtcgc      | 61    |                     |
| Db                         | 1     | atgcaataaagctctccacctgctctcttcctctgctgctcttgcgattcgtcttaagtcgc      | 60    |                     |
| QY                         | 62    | accsagaataactactcctgggtgcagtcggaactgcacatgggagcatatagcaagtgcctc     | 121   |                     |
| Db                         | 61    | accsagaataactactcctgggtgcagtcggaactgcacatggagcatatagcaagtgcctc      | 120   |                     |
| QY                         | 122   | ggttaagctgcctctgggacgcgaagaattccctccctaagtagcgcaaaatcctttccatcac    | 181   |                     |
| Db                         | 121   | gggtagcgtgcctctgggacgcgaagaattccctccctaagtagcgcaaaatcctttccatcac    | 180   |                     |
| QY                         | 182   | accctagctcgtctacaaaagaactcgtcttgcctgtagattcaacggtcttcacatttacaac    | 241   |                     |
| Db                         | 181   | accctagctcgtctacaaaagaactcgtcttgcctgtagattcaacggtcttcacatttacaac    | 240   |                     |
| QY                         | 242   | gctaaagcgaagcccaacctgcggtatgggtctgcctcagtcctcaacatccagctgcgatttat   | 301   |                     |
| Db                         | 241   | gctaaagcgaagcccaacctgcggtatgggtctgcctcagtcctcaacatccagctgcgatttat   | 300   |                     |
| QY                         | 302   | gatacagtcgtctaatcaactaaagaacatgcgtctcccatccctgcctcagctttcctcagctgtt | 361   |                     |
| Db                         | 301   | gatacagtcgtctaatcaactaaagaacatgcgtctcccatccctgcctcagctttcctcagctgtt | 360   |                     |
| QY                         | 362   | ggctatccctaccgcgaagaagctttctggggagctgcgaatgtgatgcacagccagcgaag      | 421   |                     |
| Db                         | 361   | gggtatccctaccgcggaagaagctttctggggagctgcgaatgtgatgcacagccagcgaag     | 420   |                     |
| QY                         | 422   | gagagaagaagatgataaagctctccctctgcggaagccatacatatgctctggcaggtcctg     | 481   |                     |
| Db                         | 421   | gagagaagaagatgataaagctctccctctgcggaagccatacatatgctctggcaggtcctg     | 480   |                     |
| QY                         | 482   | aaagagaatgctcccaatgctctctgacccaatgctgccttaccctacatatctttctcat       | 541   |                     |
| Db                         | 481   | aaagagaatgctcccaatgctctctgacccaatgctgccttaccctacatatctttctcat       | 540   |                     |

QY 542 gTgagccgtgtaaaagacttgaaatcagagccatcttgagccctactagtagtgaagaa 601  
Db 541 gTgagccgtgtaaaagacttgaaatcagagccatcttgagccctactagtagtgaagaa 600  
QY 602 gggagctctggccaagaaagacacagacccctgcaacaattatactacttttgcgtta 661  
Db 601 gggagctctggccaagaaagacacagacccctgcaacaattatactacttttgcgtta 660  
QY 662 tttagtgaaggaaagcttgagcactcagaaagaagaactccttgatgtagaataaggat 721  
Db 661 tttagtgaaggaaagcttgagcactcagaaagaagaactccttgatgtagaataaggat 720  
QY 722 gctgcatctgtcgagccctggccctaaatgacacagctcaatggtgtatgtaaacgctct 781  
Db 721 gctgcatctgtcgagccctggccctaaatgacacagctcaatggtgtatgtaaacgctct 780  
QY 782 ctgccaagctctgattgtagtgccaacagaaatcagtcatactgagcatgtgattggaattggcc 841  
Db 781 ctgccaagctctgattgtagtgccaacagaaatcagtcatactgagcatgtgattggaattggcc 840  
QY 842 accactcctgaagatgcatcacaatatctcctgaaggtgacacacattcttgtagaagaacat 901  
Db 841 accactcctgaagatgcatcacaatatctcctgaaggtgacacacattcttgtagaagaacat 900  
QY 902 cgcacagcgctccttggaatatcgcaccaataacttctcctactgctcaaacactcttgatg 961  
Db 901 cgcacagcgctccttggaatatcgcaccaataacttctcctactgctcaaacactcttgatg 960  
QY 962 gaccttgagacagttctactgltttgtgcatatactcttccccaacaatgagtagagaa 1021  
Db 961 gaccttgagacagttctactgltttgtgcatatactcttccccaacaatgagtagagaa 1020  
QY 1022 gcttatgtcacaagtagaagaagctgtgccagaagaaaccccaactacagatgtaaaaaatgtaa 1081  
Db 1021 gcttatgtcacaagtagaagaagctgtgccagaagaaaccccaactacagatgtaaaaaatgtaa 1080  
QY 1082 gaagcggagaagactagatgcatctactctatctgaatgagatgtgtcgaagtttgat 1141  
Db 1081 gaagcggagaagactagatgcatctactctatctgaatgagatgtgtcgaagtttgat 1140  
QY 1142 gatgcaaacctctcctctctatccaaatctgcctcgaattgccaagaagcatcctaanaact 1201  
Db 1141 gatgcaaacctctcctctctatccaaatctgcctcgaattgccaagaagcatcctaanaact 1200  
QY 1202 tgggtfaccttactgtcgtcgaaagagagagactgagacatactgctcctctatgctcctgc 1261  
Db 1201 tgggtfaccttactgtcgtcgaaagagagagactgagacatactgctcctctatgctcctgc 1260  
QY 1262 ccgcatgacagaagttataaagaatcaaatattgacaatgagccctcagcgagattggtatag 1321  
Db 1261 ccgcatgacagaagttataaagaatcaaatattgacaatgagccctcagcgagattggtatag 1320  
QY 1322 aagtaacaaaagtcggaattatgtgcatacacagatganaacctttaagactcgtgaagct 1381  
Db 1321 aagtaacaaaagtcggaattatgtgcatacacagatganaacctttaagactcgtgaagct 1380  
QY 1382 attcagcatgatacagaagactttggagcccttacttataatgggagaagtgtgagacacactg 1441  
Db 1381 attcagcatgatacagaagactttggagcccttacttataatgggagaagtgtgagacacactg 1440  
QY 1442 ttgattatattaaagaatcaagcaagcagacatataacatctacccccaagagatacact 1501  
Db 1441 ttgattatattaaagaatcaagcaagcagacatataacatctacccccaagagatacact 1500  
QY 1502 gatgtccgctcttgtatccaagagatccaaaagggtgtaanaacatltgaaagatttt 1561  
Db 1501 gatgtccgctcttgtatccaagagatccaaaagggtgtaanaacatltgaaagatttt 1560  
QY 1562 ccaattctgcagagagaataatccaataataatggaacagtgatgtgagaagaatgggcca 1621  
Db 1561 ccaattctgcagagagaataatccaataataatggaacagtgatgtgagaagaatgggcca 1620  
QY 1622 actaatcagatcctcgtgctgacccgctattactctagtttcgttaataatgagaga 1681

Db 1621 actaatcagatcctcggctgctgagccgcctattactctagtttcgttaataatgagaga 1680  
QY 1682 gatctagcttaagaactatctggccctctcctcatctgtctacaagaagaactcgtgatcaa 1741  
Db 1681 gatctagcttaagaactatctggccctctcctcatctgtctacaagaagaactcgtgatcaa 1740  
QY 1742 agaggaacccagataatgtcagacaagaagaaatgtcaaccgttttcgttatgtatgag 1801  
Db 1741 agaggaacccagataatgtcagacaagaagaaatgtcaaccgttttcgttatgtatgag 1800  
QY 1802 aaccgaagctgttacccctacagagaaatalatacaagccttctcccaatccagccgtagatg 1861  
Db 1801 aaccgaagctgttacccctacagagaaatalatacaagccttctcccaatccagccgtagatg 1860  
QY 1862 cagcttgagagatccagagttcccaagcctccacaacatcatgcaagatcaatgtgatagtt 1921  
Db 1861 cagcttgagagatccagagttcccaagcctccacaacatcatgcaagatcaatgtgatagtt 1920  
QY 1922 tttagatgttgcaagttgtcaggtttgttgatagagtgagcatactgtaattcttaagc 1981  
Db 1921 tttagatgttgcaagttgtcaggtttgttgatagagtgagcatactgtaattcttaagc 1980  
QY 1982 atggaacacagactgaactcctctctgtctctctctcctgataataccttcaaacacaa 2041  
Db 1981 atggaacacagactgaactcctctctgtctctctctcctgataataccttcaaacacaa 2040  
QY 2042 atgtctataagaagacacacatccactatcccatctcagagaagaactgtctcatgtcg 2101  
Db 2041 atgtctataagaagacacacatccactatcccatctcagagaagaactgtctcatgtcg 2100  
QY 2102 atggaaaaacccaggtctctatgattctgtgggtggtccacaactaagactttcggaacagagc 2161  
Db 2101 atggaaaaacccaggtctctatgattctgtgggtggtccacaactaagactttcggaacagagc 2160  
QY 2162 atgacccgcttactgaaggtttctcagttgtgacaagaacactgtgattatacagagac 2221  
Db 2161 atgacccgcttactgaaggtttctcagttgtgacaagaacactgtgattatacagagac 2220  
QY 2222 agttaatgaagatattccaagactctgtgagtgtaaaaaaactgcatatgaaaccaagaagc 2281  
Db 2221 agttaatgaagatattccaagactctgtgagtgtaaaaaaactgcatatgaaaccaagaagc 2280  
QY 2282 ttctccccaagaattca----- 2286  
Db 2281 ttctccccaagaattcaactatgtgattcctctgtcttgaggtaaaccaactatgtactcagata 2340  
QY 2297 ----- 2296  
Db 2341 ccaaaagaagaagtggaaatccccaagagaagtcaccaagaaaaaacagctttaagaanaaag 2400  
QY 2297 ----- 2296  
Db 2401 gataccatttgtccctgtaacgcttgtagaagaacatactgcaatagcagcaataatgag 2460  
QY 2297 -----agacccctagcactagggcaaaagca 2323  
Db 2461 ggaacaaaataagcccgaaatagaagtcacactggtgcacagcagaaggtlaagacatgaaagctg 2500  
QY 2324 tttaatgcaccccaacagctcttgaagcgcatcaacgggaataactcgtcgtactacttt 2383  
Db 2521 tgccttcaaaaccccaacagctcttgaagcgcatcaacgggaataactcgtcgtactacttt 2580  
QY 2384 cagtcagatcaagaagaaatgtgactatgtatgataccatcatcagttgtaaatgaaagaagaa 2443  
Db 2581 cagtcagatcaagaagaaatgtgactatgtatgataccatcatcagttgtaaatgaaagaagaa 2640  
QY 2444 gatttgacatttatgtatgtagatgataaatatagagcccccagcttccaagaagaaaca 2503  
Db 2641 gatttgacatttatgtatgtagatgataaatatagagcccccagcttccaagaagaaaca 2700  
QY 2504 cgacacatttatctgtcagctgagtagagagctcgggataatggagtagtagtcccca 2563

Db 2701 cgaacatttatttgcgtgagtgagaggtctctgggattatggatlgagtcctccca 2760  
QY 2564 catgttctaagaacaagggctcagagtggtcagtgctccctcagttcacaagaagtgtttc 2623  
Db 2761 catgttctaagaacaagggctcagagtggtcagtgctccctcagttcacaagaagtgtttc 2820  
QY 2624 caggaatttactgtagtgccttcttactcagcccttaccgtgggagaaactaaatgacat 2683  
Db 2821 caggaatttactgtagtgccttcttactcagcccttaccgtgggagaaactaaatgacat 2880  
QY 2684 ttggagacccttgaggccacataataagagcagaagtgtgaagaataatctctgtaacttc 2743  
Db 2881 ttggagacccttgaggccacataataagagcagaagtgtgaagaataatctctgtaacttc 2940  
QY 2744 agaaatcagggccctcgtctcccttacttcttacttctgtaacttatttctataggagaat 2803  
Db 2941 agaaatcagggccctcgtctcccttacttcttacttctgtaacttatttctataggagaat 3000  
QY 2804 cagagggcagaagcagaacctaagaaaaaacttctcagcctaataatgaaacaaacttac 2863  
Db 3001 cagagggcagaagcagaacctaagaaaaaacttctcagcctaataatgaaacaaacttac 3060  
QY 2864 ttcttgaaagtgcacaacataatggtcaccctaaagatgttgcagtgcaaggtctgag 2923  
Db 3061 ttcttgaaagtgcacaacataatggtcaccctaaagatgttgcagtgcaaggtctgag 3120  
QY 2924 gcttatttctctgtagtctgacccctggaaaaaaagtgtgacactcaggtctga ttgagccctt 2983  
Db 3121 gcttatttctctgtagtctgacccctggaaaaaaagtgtgacactcaggtctga ttgagccctt 3180  
QY 2984 ctgtgtctgcacacataacacacttgaaacctgtctcatggagacaaagtgtgacagtagagaa 3043  
Db 3181 ctgtgtctgcacacataacacacttgaaacctgtctcatggagacaaagtgtgacagtagagaa 3240  
QY 3044 ttctgtcgttttttccacatctcttgatgagcacaagaagctgtgtactctcctaataatag 3103  
Db 3241 ttctgtcgttttttccacatctcttgatgagcacaagaagctgtgtactctcctaataatag 3300  
QY 3104 gaaagaaactgcagaggtccctctgcaatatccagatggaagatcccaactttaaagagaat 3163  
Db 3301 gaaagaaactgcagaggtccctctgcaatatccagatggaagatcccaactttaaagagaat 3360  
QY 3164 tatgcttccatgcatataatggtctacataatggtatgatacctacgtgcttaataatggt 3223  
Db 3361 tatgcttccatgcatataatggtctacataatggtatgatacctacgtgcttaataatggt 3420  
QY 3224 caggaatcaagaatctgtagtgcctgcaatccagatgggagcaatgaaacatccatct 3283  
Db 3421 caggaatcaagaatctgtagtgcctgcaatccagatgggagcaatgaaacatccatct 3480  
QY 3284 attcaattcagtgagcatggtgttcaatgtaaaaaaaagagagtataaaatggtcactg 3343  
Db 3481 attcaattcagtgagcatggtgttcaatgtaaaaaaaagagagtataaaatggtcactg 3540  
QY 3344 taccaatctctatccaaggtgttttttgaagcagtggaatggtttacatccaacatggaatt 3403  
Db 3541 taccaatctctatccaaggtgttttttgaagcagtggaatggtttacatccaacatggaatt 3600  
QY 3404 tggcgggtggaatgacctatctgagcagcatctacatgctggagatgagcaccctttctg 3463  
Db 3601 tggcgggtggaatgacctatctgagcagcatctacatgctggagatgagcaccctttctg 3660  
QY 3464 gttgtaacagcaatagtgctcagaactccctgggaatggtctctggaacataatgagaattt 3523  
Db 3661 gttgtaacagcaatagtgctcagaactccctgggaatggtctctggaacataatgagaattt 3720  
QY 3524 cagattacagcttcaggaacaataatgagacagtggtgcccagaagctggccagattcatat 3583  
Db 3721 cagattacagcttcaggaacaataatgagacagtggtgcccagaagctggccagattcatat 3780  
QY 3584 tcaggaatcaaatgacctgagcacaagaaggtccctttcttgatcaagtgagctg 3643  
Db 3781 tcaggaatcaaatgacctgagcacaagaaggtccctttcttgatcaagtgagctg 3840

QY 3644 ttggaccaaatgattatctcagggcatcaagaccaggtgtcccgctcagaagttctccagc 3703  
Db 3841 ttggaccaaatgattatctcagggcatcaagaccaggtgtcccgctcagaagttctccagc 3900  
QY 3704 cttcaatctccaggtttatcatcatgtaatgtaattcttgatgggaagaagtgtgacattat 3763  
Db 3901 cttcaatctccaggtttatcatcatgtaatgtaattcttgatgggaagaagtgtgacattat 3960  
QY 3764 cgaggaaattccacitggaaccttaatggtctctcttggcaaatgtagtcatctgggata 3823  
Db 3961 cgaggaaattccacitggaaccttaatggtctctcttggcaaatgtagtcatctgggata 4020  
QY 3824 aaacaatatctttaaaccctccaattatgtctcgatatacatccgtttgacccaactcat 3883  
Db 4021 aaacaatatctttaaaccctccaattatgtctcgatatacatccgtttgacccaactcat 4080  
QY 3884 tataagattcagcagcaactcttcgcatgtagtctgtagtgggtgtgatttaattagtgtcagc 3943  
Db 4081 tataagattcagcagcaactcttcgcatgtagtctgtagtgggtgtgatttaattagtgtcagc 4140  
QY 3944 atgcatgtggagatggaggttaaagcaatatcagatgacagattactgcttcatctac 4003  
Db 4141 atgcatgtggagatggaggttaaagcaatatcagatgacagattactgcttcatctac 4200  
QY 4004 tttaaccaatagtttggccacctgtctctctcaaaaagctcgacttcaacctccaaggagag 4063  
Db 4201 tttaaccaatagtttggccacctgtctctctcaaaaagctcgacttcaacctccaaggagag 4260  
QY 4064 agtaagctctggagagacctatgtagtaataatccaagaagtgtctgcaagtgtactccag 4123  
Db 4261 agtaagctctggagagacctatgtagtaataatccaagaagtgtctgcaagtgtactccag 4320  
QY 4124 aagacatgaaagtccaagagttaactctacaggtgaggtaaactctctcttaccagcatg 4183  
Db 4321 aagacatgaaagtccaagagttaactctacaggtgaggtaaactctctcttaccagcatg 4380  
QY 4184 tatgtgaagagttccctcatctccagcagttcaagaatggccatcagtgactctctttt 4243  
Db 4381 tatgtgaagagttccctcatctccagcagttcaagaatggccatcagtgactctctttt 4440  
QY 4244 cagaatggcaaatgtaagttttttaggggaaatcagaagctctcttcaaccctgtgtgac 4303  
Db 4441 cagaatggcaaatgtaagttttttaggggaaatcagaagctctcttcaaccctgtgtgac 4500  
QY 4304 tctctagaccacccgttactgactcgtcaactctgcaattcaacccccaagatgggtgac 4363  
Db 4501 tctctagaccacccgttactgactcgtcaactctgcaattcaacccccaagatgggtgac 4560  
QY 4364 cagattgccttgagatgtaggttcttgggtctgagagcacaagaccttactga 4417  
Db 4561 cagattgccttgagatgtaggttcttgggtctgagagcacaagaccttactga 4614

Search completed: November 18, 2001, 05:14:47  
Job time: 36419 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2001, 01:19:45 ; Search time 22045.3 Seconds  
(without alignments)  
3099.125 Million cell updates/sec

Title: US-09-689-430-1\_COPY\_419\_4835  
Perfect score: 4417  
Sequence: 1 catgaatagactctcca.....agcacagagactctactga 4417

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba1:\*  
2: gb\_ba2:\*  
3: gb\_ba3:\*  
4: gb\_in1:\*  
5: gb\_in2:\*  
6: gb\_in3:\*  
7: gb\_om:\*  
8: gb\_ov:\*  
9: gb\_pat1:\*  
10: gb\_pat2:\*  
11: gb\_ph:\*  
12: gb\_pil:\*  
13: gb\_p12:\*  
14: gb\_p13:\*  
15: gb\_p14:\*  
16: em\_ba1:\*  
17: em\_ba2:\*  
18: em\_fun:\*  
19: em\_htgo\_hum:\*  
20: em\_htgo\_inv:\*  
21: em\_htgo\_rod:\*  
22: em\_htg\_hum1:\*  
23: em\_htg\_hum2:\*  
24: em\_htg\_hum3:\*  
25: em\_htg\_hum4:\*  
26: em\_htg\_hum5:\*  
27: em\_htg\_hum6:\*  
28: em\_htg\_hum7:\*  
29: em\_htg\_hum8:\*  
30: em\_htg\_inv1:\*  
31: em\_htg\_inv2:\*  
32: em\_htg\_other:\*  
33: em\_htg\_rod:\*  
34: em\_hum1:\*  
35: em\_hum2:\*  
36: em\_hum3:\*  
37: em\_hum4:\*  
38: em\_hum5:\*  
39: em\_hum6:\*  
40: em\_hum7:\*  
41: em\_in:\*  
42: em\_om:\*  
43: em\_or:\*

44: em\_ov:\*  
45: em\_pat:\*  
46: em\_ph:\*  
47: em\_pl:\*  
48: em\_ro:\*  
49: em\_sts:\*  
50: em\_sy:\*  
51: em\_un:\*  
52: em\_v1:\*  
53: gb\_sts1:\*  
54: gb\_sts2:\*  
55: gb\_sts3:\*  
56: gb\_sy:\*  
57: gb\_un:\*  
58: gb\_v11:\*  
59: gb\_v12:\*  
60: gb\_htg1:\*  
61: gb\_htg2:\*  
62: gb\_htg3:\*  
63: gb\_htg4:\*  
64: gb\_htg5:\*  
65: gb\_htg6:\*  
66: gb\_htg7:\*  
67: gb\_htg8:\*  
68: gb\_htg9:\*  
69: gb\_htg10:\*  
70: gb\_htg11:\*  
71: gb\_htg12:\*  
72: gb\_htg13:\*  
73: gb\_htg14:\*  
74: gb\_htg15:\*  
75: gb\_htg16:\*  
76: gb\_htg17:\*  
77: gb\_htg18:\*  
78: gb\_htg19:\*  
79: gb\_htg20:\*  
80: gb\_htg21:\*  
81: gb\_htg22:\*  
82: gb\_htg23:\*  
83: gb\_htg24:\*  
84: gb\_htg25:\*  
85: gb\_pr1:\*  
86: gb\_pr2:\*  
87: gb\_pr3:\*  
88: gb\_pr4:\*  
89: gb\_pr5:\*  
90: gb\_pr6:\*  
91: gb\_pr7:\*  
92: gb\_pr8:\*  
93: gb\_pr9:\*  
94: gb\_ro1:\*  
95: gb\_ro2:\*  
96: gb\_in4:\*  
97: gb\_pr10:\*  
98: em\_ba3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Query Match | Length | DB | ID       | Description       |
|------------|--------|-------------|--------|----|----------|-------------------|
| 1          | 4407.4 | 99.8        | 9354   | 9  | AR003585 | AR003585 Sequence |
| 2          | 4357.4 | 98.7        | 4670   | 9  | AR110040 | AR110040 Sequence |
| 3          | 4148.8 | 93.9        | 4278   | 10 | I08644   | I08644 Sequence 4 |
| 4          | 4147.2 | 93.9        | 4281   | 10 | I08643   | I08643 Sequence 3 |
| 5          | 4085   | 92.5        | 4348   | 10 | I08642   | I08642 Sequence 2 |
| 6          | 4085   | 92.5        | 4551   | 10 | I08641   | I08641 Sequence 1 |
| 7          | 3743.8 | 84.8        | 5035   | 9  | AR034084 | AR034084 Sequence |
| 8          | 3743.8 | 84.8        | 5035   | 9  | AR071306 | AR071306 Sequence |



|    |      |   |      |
|----|------|---|------|
| Qy | 61   | caccagaagaatactactggtggtgcaggggaacgtgcataatgagactatgcaaaagtgcact    | 120  |
| Dy | 3024 | CACCAGAGATACTACTGGTGTGACAGTGAAGCTGTATGGAGCTATATGCAAGATGATCT         | 3083 |
| Qy | 121  | cgtgtagctbctcgtgtgagcgcgaagatctctccctagagtgcgaanaatctttccatcaa      | 180  |
| Dy | 3084 | CGGTGACCTCCGTGTGAGCCAGAGATTTCCTCTCTCAAGTGCAGAAATCTTTTCCATTCAA       | 3143 |
| Qy | 181  | caacctagtcgtgtacaaaagaactcgtttgttagaattcaacgtgtcaacctttcaacat       | 240  |
| Dy | 3144 | CACCTCAGTGTGTACAAAAGACTCTGTTGTTGATGATTCAGGCTTCACCTTTTCACAT          | 3203 |
| Qy | 241  | cgtctaaagccaaagccaaacctcgatgggtctgtctaggtctctaccatccaagctcgagttta   | 300  |
| Dy | 3204 | CGCTAACCCAGAGCCACCCTGGATGGGTGCTGTACAGTCCATCCAGGCTGAGGTTTA           | 3263 |
| Qy | 301  | tgatataagtgatcatatcaacttaagaacaatggtctcccatctctgcagcttctcaatgcgt    | 360  |
| Dy | 3264 | TGATACAGTGTGTATTACATTAGAACATGAGCTTCCACTCTGTGAGTCTTTCATGCTGT         | 3323 |
| Qy | 361  | tgggttactctacttggaaagctctctgaaggagctgataatgataatgataatgcacgaatgc    | 420  |
| Dy | 3324 | TGGTGTACTCTACTGTGGAAGCTTCTGAGGAGCTGAAATGATGATGATCAAGACAGTCAAG       | 3383 |
| Qy | 421  | ggagaagaagaatgataaagctctctccctggtgtgaaagccatacatatgtctggcaggtct     | 480  |
| Dy | 3384 | GGAGAAAGAGATGATTAAGATCTTCCCTGGTGGAAAGCCATACATATGCTGTGGCAGGTCT       | 3443 |
| Qy | 481  | gaaagagaatgtgccaatgtgcctctaacacacgtgccttaccatctatctcttcca           | 540  |
| Dy | 3444 | GAAAGAGAAATGGTCCATATGGCCCTTACACCACAGTGTCTCTTACTCTCATATATCTTTTCA     | 3503 |
| Qy | 541  | tgtgagacctgtgataaagaacttgatattcaagcctcatcttgaaagcccaactagatgtgaaaga | 600  |
| Dy | 3504 | TGTGAGACCTGTGTAAGACCTTGAAATTCAGGCTCATATGGAAGCCCTACTAGTATGTAAGGA     | 3563 |
| Qy | 601  | aggagagcttggccaaggaagaacacagaacctgcacaaattatatactacttttgcgt         | 660  |

|    |      |   |      |
|----|------|---|------|
| Db | 3624 | ATTGTGAAGAAAGGAAAAAGTTGGAGCATCAACAACAAAGAACTCCTTGATGACAGATAGGGA       | 3683 |
| Qy | 721  | tgctgcagctcgtgcctcgggcccctgggctcaaaatgcaacaagtcgaatgtgtatgtaaaaggtc   | 780  |
| Db | 3684 | TGCTGCATTCCTGCTCGGGCCCTGGCTTAATATGCACACAGTCGAATGGTTATGTRAAACGGTC      | 3743 |
| Qy | 781  | ctctgcacgctctgatttggatgcatgcacacagaaatcagtcatttggcatgltgatttgaaatggtg | 840  |
| Db | 3744 | TCTGCCAGGCTGTATGGAGTGCACACAGAAATCAGTCCTTTGGCATGTGATGTGAATGGG          | 3803 |
| Qy | 841  | caccacaccttgaaagtgcactccaatatctctgtaaggtcaccaattcttctgtgagaacaa       | 900  |
| Db | 3804 | CACCACTCCGGAAGTGACATCAATTAATTCCTCGAAGGTCAACAAATTTCTGTGAGGAACCA        | 3863 |
| Qy | 901  | tcgcacagcgtctccttggaaatctcgcacaaatacttcctactctgtcacaacactcttgat       | 960  |
| Db | 3864 | TTCGCACGGCTCCTTGGAATCTCGCCAACTTAACCTTCCTTACTGGTCAACACACTTGTAT         | 3923 |
| Qy | 961  | ggacacttgacagttcttactglttctgtcatatctctcccaaccaatgatgacatgga           | 1020 |
| Db | 3924 | GGACCTTGGACAGTTCTACTGTTTGTATATCTCTCCACCAACAGATGAGCATGGA               | 3983 |

Db 3984 AGCTATGTCAAAGTAGACAGCTGTCACAGAGAACCCCACTACGATGAATAATATGCA 4043

Db 404 AGAGCGAAGACTATGATGATCTTACTGATTTCTGAATGGATGTGGTCAAGTTTCA 4103

Db 4104 TGAAGACAACTCTCTCTTATCCAAATTCGCTAGTTSCCAAGAACATCTAAAC 4163  
OY 1201 ttgggtacataactctgtctgtgaagagagactgggactatgctccctagtccctgc 1260  
Db 4164 TTGGGTACATTATTCGTCTGGAAGAGAGACTGGACATATCTCTCTAGTCCCTGC 4223  
OY 1261 ccccgatgacagaagtataaagaatcaatattgacaatggcctcaagcgatggttag 1320  
Db 4224 CCCCAGTACAGAAAGTTTAAAGTCAATATTTAACATAGGCCCTCAGCCGATGGTAG 4263  
OY 1321 gaagtacaaaaaagtcgcattatgcatatagacagatgaacacttgaagctgtgaagc 1380  
Db 4284 GAAGTACAAAAAGTCCGATTATGGCATACAGATGAACCTTTAAGACTCCTGAAGC 4343  
OY 1381 tattcagcatgaatcagaatcttgggaacttactcttaagggaagttgagacacact 1440  
Db 4344 TATTACATGAATCAGAAATCTTGGACCTTTACTTTATGGGAGATGGACACACT 4403  
OY 1441 gtgattatatttaagaatcaagaacagacacataacactacccctcaaggaatcac 1500  
Db 4404 GTTATTATATTATTAAGATACAGACAGACATATTAACCTTAACCTCAGGAAATCAC 4463  
OY 1501 tgatgtccgtcccttgtatccaagagatlaacaaaagtgtaaaaacatttgaagattt 1560  
Db 4464 TGATGTCCGTCTTGTATTCAGAGAGATTACAAAAGGTGTAAACATTTGAAGGATTT 4523  
OY 1561 tccaactctgcagagagaataatacaataatgacagtgacgtgagaagatggcc 1620  
Db 4524 TCCATTCTCGCAGAGAAATATTCAAATTAATTAATGACAGTGTAGAAAGATGGGCC 4583  
OY 1621 aactaaatcagaatcctcgtggtgcctgcagccgtatctactgaagttgttgaataggag 1680  
Db 4584 AACATAATCAGATCTCTGGTCCCTGACCCGCTATTACTCTAGTTTGGTTAATGAGAG 4643  
OY 1681 agatcagctcagagactatggccctccctcatctgctacaaagaatctgtagatca 1740  
Db 4644 AGATCTAGCTTCAGAGACTCATTTGGCCCTCTCCATCTGTACAAAAGATCTGAGATCA 4703  
OY 1741 aaggggaaacagaataatgtgcagaagaagaatgcatcctgttctgtatattgtga 1800  
Db 4704 AAGGGAAACCAATATATGTGCAGCAAGAGAAAGTATCATCTGTCTGTATTTGATGA 4763  
OY 1801 gaaccgaagctgtgacctcacaagaatatacaacgcttctccccaatccagctgaagt 1860  
Db 4764 GAACCGAAGCTGTACTCTCACAGAGATATATACAGCGTTTCTCCCAATCCAGCTGAGT 4823  
OY 1861 gcagcttgagaatccagaatcccaagcctcccaatcatgcaacagcatcaatggctatgt 1920  
Db 4824 GCACCTTGAGGATCCAGAGTTCCCAAGCCTCCACATCATGCACAGCATGAGCTATGT 4883  
OY 1921 tttagatagtttagagttgtcgaagttgttgatgaagtgagcatgaactgtaattcgaag 1980  
Db 4884 TTTTGTATGATTTGCAATTTGTGATTTGTTCATGAGGTGGCAATGATCATTTCTTAAG 4943  
OY 1981 catggaagcaagaactgaactccttctgtctctctcctgatatatacctcaaacaa 2040  
Db 4944 CATTTGAGACACAGCTACTCTCTTCTGTCTCTCTCTGTGATATACCTTTCAAAACAA 5003  
OY 2041 aatgtctatgaagacacactcaacctatcccatctcagagagaacatgcttcatgctc 2100  
Db 5004 AATGTCTATGAAGACACACTCACCTATTCCTCATCTCAGAGAAACTGCTTCATGTCTC 5063  
OY 2101 gatggaaaacccaggtctatgagttctggtgtgcacaactcgaacttgggaacagag 2160  
Db 5064 GATGGAAAACCCAGGTTATGAGATTGTGGGTGGCCACAACCTTTTGGGAACAAGG 5123  
OY 2161 catgaccgctctactgaaggttctctagtgtgacaagaacacgtgtattatitacagga 2220  
Db 5124 CATGACCGCCTTACTGAAGTTTCTAGTTGTGACAAAGAACACTGTGTATTTATTCGAGGA 5183  
OY 2221 cagttatgaagataatccagcatactgtctgagtaaaaacatgcatctgaaccaagaag 2280

Db 5184 CAGTTATGAAGATATTTCAGCATCTTGTGAGTAAAAAACAAATGCCATTGAACCAAGAG 5243  
OY 2281 ctctcccaagattcaagaccctctagactaggcaaaagaatttatgtcaccaccac 2340  
Db 5244 CTCTCCCAAGAAATTCAGACACCCCTAGCATGAGCAAAAAGCAATTTAATGTGCCACCCAC 5303  
OY 2341 agcttgaagccatcaagcggaataactcgtactactctctcaagtaagtaagga 2400  
Db 5304 AGCTTGAAACGCCATCACAGCGGAATTAACCTGTACTCTCTTGTAGTCAGATCAAGAGGA 5363  
OY 2401 aatgacbatgataacacatacatgaattgaaatgaagaagaagatttgaacattatga 2460  
Db 5364 AATTGACTATGATGATACCATATCATGTTGAATGAAGAGAAAGATTTTGACATTTATGA 5423  
OY 2461 tgaagatgaaaaatcaagggccccgcagcttccaagaagaacacgaactatctatgtc 2520  
Db 5424 TGAGAGATGAATAATCAGAGCCCGCCAGCTTTCAAAAGAAAACAGACATATTTATTTGC 5483  
OY 2521 tgcagtgagaaggtctcgggataatgagatgagtagtgcctcccaatgttctaaagaacag 2580  
Db 5484 TGCAGTGGAGAGGCTCTGGGATTTATGGGATGATGACTCCCCACATGTTCTTAAAGAACAG 5543  
OY 2581 ggtcagagtggtcagtgctccctcagttccaagaagaagttgttccaaggaattactatg 2640  
Db 5544 GGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTTCCAGGAATTTACTGATGG 5603  
OY 2641 ctctcttaactcagcccttatccgttgagagaactaaatgaacatttgggactcctggggcc 2700  
Db 5604 CTCTTACTCAGCCCTTATACCTGTGAGAACTTAATGAACATTTGGACACTCGTGGGCC 5663  
OY 2701 atataaagcagaagttgagaataatcatgttaacttcaagaaatcagagctctcg 2760  
Db 5664 ATATATTAAGACCAAGTGTGAAGTAATATCATGTGTAATCTTCAGAAATCAGGCTCTCG 5723  
OY 2761 tccctactcctctatctcagcttattctctatgagaagaatcaagggcaagaga 2820  
Db 5724 TCCCTATTTCTTATCTTACCTTATTTATTTATGAGGAAGATCAGAGGCAAGGACAGA 5783  
OY 2821 accatgaaaaaacttgcagccctaatagaacaaacttacttggaaagtgcaca 2880  
Db 5784 ACCTAGAAAACCTTTGTCAAGCCTTAATGAACCAAAACTTACTTTGGAAAGTGCAACA 5843  
OY 2881 tcaatgtgcaacccaataagaatgagttgactgcaagccctgggcttatttctctgagt 2940  
Db 5844 TCATATGGACACCCTAAAGATGAGTTGACTGCAAAAGCCTGGGCTTATTTCTGTGATGT 5903  
OY 2941 tgaccctggaaaaagaatgtgcaccgaagcgtgaattggagccctctgtgtcgcacactaa 3000  
Db 5904 TGACCTGGAAAAGATGTGCACCTCAGGCTGATTTGACCCCTTGTGTGCTGCCACACTAA 5963  
OY 3001 cacactgaacccctgtcatatgggaagaagaagtgacagtgtaaggaatttctgtgtttcac 3060  
Db 5964 CACACTGAACCTCTCATGTGGAGACAAAGTACAGTACAGGAATTTCTCTGTTTTCAC 6023  
OY 3061 catcttggatgagaccaaactgttacttcaactggaataatggaagaacatgcaagc 3120  
Db 6024 CATTTTGTATGAGACCAAAACCTGTACTTCTACATGAAATATGGAAGAATCTCAGGCGC 6083  
OY 3121 tccctgcaatatccagaatggaagatcccaactttaaagagaatatcgtctcatalgcaat 3180  
Db 6084 TCCCTGCAATATCCAGATGGAAGATCCACCTTTTAAAGAGAAATATGTGCTTCCATGCAAT 6143  
OY 3181 caatggtctacataatgatacacactacgtgtgtatgagtgtaagatcaaaagattcg 3240  
Db 6144 CAATGGCTACTAATATGATACACTACCTGGCTTATGATGCTCAGAGTCAAAAGGATTCG 6203  
OY 3241 atggtatctgtcagcatggtgcaagcaatgaaaaacatcatctatcaattcagtgagca 3300  
Db 6204 ATGTATCTCTGCTGCAATGGGCAACCAATGAAGAAATCATCTATTTCAATTCAGTGGACA 6263  
OY 3301 tgtgttcaactgtacgaaaaaagaagagatbaaaaatggcaactgtatcaactctatccagg 3360  
Db 6264 TGTGTCTACTGTACGAAAAAAGAGAGATATTAATATGGCACTGTACATCTCTATCCAGG 6323

3361 tctttttgagacagtggaatgtlaccatccaaagctggaaatttggcgggtggaatgcct 3420  
|||||  
6524 TGTTTTGAAGACAGTGAATGTTACATCCAAAGCTGGAATTTGGCGGTGAATGCCT 6383  
3421 tattggcgaacatccacatgctggatgagacacttcttcggtgtacagcaataatgct 3480  
|||||  
6384 TATTGGCAGCATCTACATGCTGGGATGAGCAGCTTTTTCGGTGTACAGCAATTAAGC 6443  
3481 tcagactccctgggaatgagcttcggaacacatagagaattcagattacagcttcagg 3540  
|||||  
6444 TCAGACTCCCTGGGAATGGCTTCTGGACACATTAGAGATTTTTCAGATTACAGCTTCAGG 6503  
3541 acaatgatgacagtgaggcccaaaagctggccagacttcatttcggaatcaatcaatgc 3600  
|||||  
6504 ACAATATGACAGTGGGCCCAAGCTGGCCAGACTTCAATTATTCGGATCATCATGC 6563  
3601 ctggaagcaacgaagacccttcttcttgatcaaggtgagctgctgtggcaacatgatat 3660  
|||||  
6564 CTGGAAGCAACGAAGACCCCTTTCTTGGATCAAGGTGAGTCTGTGGCACCAATGATATAT 6623  
3661 tcacggacatcaagaccagggtgctccatagagaattcaccagcttcacatctcagct 3720  
|||||  
6624 TCACGGCATCAAGACCCAGGGTGCCGTCAGAAATTCTCCAGCCTCTACATCTCTCAGTT 6683  
3721 tatcatcatgataagcttcttgatgggaagaagtgagacactlccgagaatctccactgg 3780  
|||||  
6684 TATCATCATGATAGTCTTTGATGGAGAGATGGCAGACTTATTCAGGAATTTCCACTGG 6743  
3781 aaccttaatgcttcttcttgcaaatgttgatcattcgtggaataaaacacataattttaa 3840  
|||||  
6744 AACCTTAATGGCTCTTGGCAATGTGATTCATCTGGGAATAAACATAATATTTTAA 6803  
3841 cccccaattatgtctgatatacatccgtttgacccaactcattatagatcttcgagcac 3900  
|||||  
6804 CCTTCGCAATTTATGCTCATACATCCGTTTGCACCCAACTCATATAGCATTCGCGAGCC 6863  
3901 tcttcgcaatgagatgagtgagcttgatattaaatagttgcaagatgccaatgggaatgga 3960  
|||||  
6864 TCTTCGCAATGAGTGTGATGGGCTGTGATTTAAATAGTTGCACATGCCATGGGAATGGA 6923  
3961 gaggtaagaacatatacagatgacagatactgcttcacatctcattaccatattgttgc 4020  
|||||  
6924 GAGTAAGCAATATCAGATGACAGATTAATCTCATCTTACCAATATGTTTTGC 6983  
4021 caactggtctcttcaaaagctcgagctccaactccaaggaaggaatgagcttcgagacc 4080  
|||||  
6984 CACCTGGCTCTCTTCAAAAGCTCGACTTCACCTCCAAGGAGAGATTAATGCTTGAGACC 7043  
4081 tcaagtgaaataatccaaagaagtgagcttgcaatgagacttcagaagaacaaatgaagtcac 4140  
|||||  
7044 TCAGGTGAATATATCCAAAGAGTGGCTGCAAGTGAGACTTCAGAAACAAATGAAGTAC 7103  
4141 agggatgaactatccaggaagtaaaatctctgttaccagcatgatatgtgaaggaattcct 4200  
|||||  
7104 AGGAGTAATCTACTAGGAGTAATAATCTCTTACACAGCATGTATGTAAAGAGTTCTCT 7163  
4201 catctccagcagtcgaagtggccatcagctgagctctcttcttccagaatggcaaatgaa 4260  
|||||  
7164 CATCTCCAGAGTCAAGATGGGCATCAGTGGACTCTCTTTTTCAGAAATGGCAAACTAA 7223  
4261 ggttttccaagggaatcaagactcctcaacaccgtgtgtgaaactccttagaccacgctt 4320  
|||||  
7224 GGTTTTTCAGGGAATCAAGACTCTTCAACACTGTGGAATCTCTTGAAGACCCACCGTT 7283  
4321 actgactcgtcactcttgaaattcaacccccagagtggtgtgacacagatctgccttagagt 4380  
|||||  
7284 ACTGACTGCTAACCTTCAATTCACCCACAGTTGGGTGACACAGATTGGCCCTGAGGAT 7343  
4381 ggaagttctggcgtgagagcaacagactctactga 4417  
|||||  
7344 GGAAGTTCTGGGCTGCGAGGCACAGGACCTTACTGA 7380

RESULT 2  
AR110040 4670 bp DNA PAT 14-FEB-2001  
LOCUS AR110040 Sequence 41 from patent US 6114148  
DEFINITION  
ACCESSION AR110040  
VERSION AR110040.1 GI:12826316  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE  
AUTHORS 1 (bases 1 to 4670)  
TITLE Seed, B. and Haas, J.  
JOURNAL High level expression of proteins  
Patent: US 6114148-A 41 05-SEP-2000;  
FEATURES  
source 1..4670  
BASE COUNT 1325 a 1072 c 1036 g 1237 t  
ORIGIN  
Query Match 98.7%; Score 4357.4; DB 9; Length 4670;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4358; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
59 gccacacagaagatactaccctggtgagtgagacgtcatgtggaactatataagatgat 118  
84 gccaccagaaagatfactactgctggatgagacgtcatgtggaactatataagatgat 143  
119 ctcggtgagctgctgtgtgagcgaagattccctcctagagtgcgaatcttccatc 178  
144 CTCGGGAGACTCCCTGTGACGACCAAGATTCTCTCTAAGAGCCAAATCTTTTCCATTTC 203  
179 aacactcagctgtgtataaaagactcgtgtgagaatcaagctcaccttcaac 238  
204 AACACCTCAGTGTGTACAAAAGACTGTTGTGAAATTCAGGATCACGATTCACCTTTTCAAC 263  
239 atcgcataacccaagccaccctcgtgagctgctgtagctccatccatccagctgagtt 298  
264 ATCGCTAACCCAAAGGACACCCTGGATGGGTGCTGCTAGGTCTACCATTCAGAGGTT 323  
299 tatgtacagtggtcattacacttaagaacatgcttccatccctgtacatgctcctatgct 358  
324 TATGATACAGTGTGATTAACATTAAGAAACATGCGTCCATCCATGCTGATCATGCT 383  
359 gttggtatcctacttgaaagactctcgaaggagtgatgatgatcagaacgaacaa 418  
384 GTTGTGTATCTTACTTGAAAGCTTCTGAGGAGCTGAATATGATGATCAGACAGTCAA 443  
419 agggagaagaagatgataaagcttccctgtgtggaagccatataatgcttcgagctc 478  
444 AGGGAAGAAAGATGATTAAGTCTTCCCTGGTGAACCATPACATATGCTTGAGAGTGC 503  
479 ctgaaagaagaatgctccaaatgagctctgaaacacatgagcttgcttaccatcatccttc 538  
504 CTGAAGAAGAAATGTCATGACATGGCTCTGACACCATGCTGCTTACTCATATATCTTCT 563  
539 catgtggacctgttaaagaacttgaaatccagcctccatctggagccctactagatgaga 598  
564 CATGTGACCTGTGTAAAGACTTGATTCAGGCTCATGTGAGCCCTACTGATGTATGAGA 623  
599 gaaaggagctctggccaaggaagaacacagacacttgacacaaattatatactacttctgc 658  
624 GAAAGGAGTCTGGCCAAAGAAAGACACAGACACTTGCAAAATTTATATCTACTTTTGGCT 683  
659 gtaattgatgaagggaagaagttgacatccagaacaagaactcctgtatgacagatag 718  
684 GTATTGTGATGAAGGAAAAAGTTGGCAGTCAAGAAAGAAAGAACTCTGATGATCAGAGTAGG 743  
719 gatgtcatcctgctcgggctggcctaaatgaacacagatgataatgtaataagagg 778  
744 GATGCTGATCTGCTGGGCTGCGGCTTAATAATGACACAGTCAATGATTATGTAACAGG 803

QY 779 tctctgcaggtctgattgattccacaggaaatcaatctatctggaatggtgattgaaatg 838  
|||||  
Db 804 TCTCTGCCAGGTCTGATTGGATGCCACAGAAATCACTATTGGCATGTGATTGGAAATG 863  
QY 839 ggcacacactccttaagtcaatcaataltcctcgaaggtgcacacattctctgtgaagaa 898  
|||||  
Db 864 GGACACCACTCCTGAAGTGCATATTCCTCGAAGGTGCACACATTTCTGTGTAGGAAC 923  
QY 899 catcgccaggcgctcctgtgaatctcgccaataacttcccttactgtctcaaacctcttg 958  
|||||  
Db 924 CATCGCCAGCGCTCTTGGAATCTGCCAATTACTTCTCTACTGCTCAACACTCTTGG 983  
QY 959 atggaccttgagagatttctactgitttjctcatatctcttccacccaatgtatgcatg 1018  
|||||  
Db 984 ATGGACCTTGACAGTTCCTTACTGTTTGTCTATCTCTTCCACCACTGATGGCATG 1043  
QY 1019 gaagcttatgtlcaaaatagaacgtgtccagaggaaacccaactacgaatgaanaaat 1078  
|||||  
Db 1044 GAAGCTTATGTCAAAGTAGACAGCTGTCCAGAGAACCCCACTACGAATGAANAATAT 1103  
QY 1079 gaagaagcggaaagactatgattgattcttactgtctgtgaatgtgtgtcaggttc 1138  
|||||  
Db 1104 GAAGAACCGGAAGACTATGATGATGATCTTACGATTCTGAAATGATGTGTCAGCTTT 1163  
QY 1139 gatgatgacaacactcctcctcttaltccaaatcgcgtcagttgcgaagaagcatctctaa 1198  
|||||  
Db 1164 GATGATGACAACTCTCTTCTTATCCAAATTCGCTCACTTGGCCAGAAAGCATCTCTAAA 1223  
QY 1199 acttgggtacattacattgtctgtcgaaggaggactggactatgtctccttactgtctc 1258  
|||||  
Db 1224 ACTTGGGTACATTACATTCGTGCTGAAGAGAGAGACTGGGACTATGCTCCCTTAGTCTC 1283  
QY 1259 gcccccgaatgaacgaagttataaaatgaatattgaacaatggccctcgaagcggattggt 1318  
|||||  
Db 1284 GCCCCCCGATACAGAAAGTTATAAAGCAATATTTGAAACATGATGCCCTCAGCGGATTTG 1343  
QY 1319 aggaagatcaaaaaaagtcgaattatgtgcaatacagaatgaacacttaagactcgtgaa 1378  
|||||  
Db 1344 AGGAAGTACAAAAAGTCCATTTATGTGCAACGATGAATAAACCTTTAAGACTCGGAA 1403  
QY 1379 gctattcaagatgaatcaagaatcttgggaacttacttctatgtgggaagtgtgagacaa 1438  
|||||  
Db 1404 GCTATTGAGCATGATCAGGAATCTTGGGACCTTTACTTTATGGGGAAGTTGGAGACACA 1463  
QY 1439 ctgttgaattatlttaagaatcaagaacagacacataatacatctacccctcaggaatc 1498  
|||||  
Db 1464 CTGTTGATTATATTTAAGAAATCAAGCAAGCAACATATACATCTACCTCAGGAATC 1523  
QY 1499 actgatgtcgttccttctgtatltcaagaagatltacaaaaggtgtlaaaacaattggaagat 1558  
|||||  
Db 1524 ACTGATGTCCGCTCTTGTATTCAGAGGAGATTACCAAAAGGTGTAAANACTTTGAAAGGAT 1583  
QY 1559 ttctcaatltctgcaggagaataataltcaaatataaatgaacagtgagatgagatgag 1618  
|||||  
Db 1584 TTTCCAATTCGTGCAGAGAAATATTCAAATATTAATGAGACAGTGTGAAGAGATGGG 1643  
QY 1619 ccaactaatatcagatctcggtgcctggaacccgctataactctagttcgttlaataatgag 1678  
|||||  
Db 1644 CCAACTTAATCAATCTCGTGCCTGACCCGCTATTAATCTAGTTCTGTTAATATAGAG 1703  
QY 1679 aagagatcttcaagagactatgtgacctctcctctatctgcttcaaaaagatctgtgat 1738  
|||||  
Db 1704 AGAGATCTAGCTTCAGACATCTATTGGCCCTCTCTCTCATCTGCTACAAAGAAATCTGAT 1763  
QY 1739 caaagagaaacagaaatagtctcagaacaaggaaatgtcaatccgttcttctgattatgat 1798  
|||||  
Db 1764 CAAGAGAGAAACAGATATATGTCAAGACAGAGAAATGTCATCGTGTCTGTGATTTGAT 1823  
QY 1799 gagaacccaagctgtgtactcacaagaataatacaacgcttcttcccaatccaagctgga 1858  
|||||  
Db 1824 GAGAACCGAAGCTGTGTAACCTACAGAGAAATATCAACAGCTTTCTCCCAATCCAGCTGGA 1883

QY 1859 gtgcagcttgagatccagagttlccaaagcctccacaatcagcaacatcaatggtcat 1918  
|||||  
Db 1884 GTGAGCTTGAGATTCAGAGTTCCAGAGTTCCAGCTCCAACTCATGTGCACACATCAATGCGTAT 1943  
QY 1919 gttttgatatgttgcagttgtcagttgttlttgcaatgaagtgagcatctgttatctta 1978  
|||||  
Db 1944 GTTTTATGATGTTTGCAGTTGTGCTAGTTGTTTGTGATGAGGTGCATGCTGATCATTTCTA 2003  
QY 1979 agcaattgagacagagactgacttcccttctgtctctctctctctgtatataaccttaaacac 2038  
|||||  
Db 2004 AGCATTTGGACACAGACTGACTTCCTTTCGTCTTCTTCTGTGATTTACTCTCAAAAC 2063  
QY 2039 aaaaatgtctatgaagacacactacactatcccatctccaagaagaacgtgtctatg 2098  
|||||  
Db 2064 AAAATGCTATGAGACACACTACCTATTCCTCATTTCCAGAGAAACTGTCTTTATG 2123  
QY 2099 tcatgtgaaaccccaagttctatgtatctgtggtgtgcacaaactcgaacttctggaacaga 2158  
|||||  
Db 2124 TCGATGAAACCCAGSTCTATGATTCGTGGGGTCCCAACTCAGACTTTCGGAACAGA 2183  
QY 2159 ggcatagcgccttactcaggaaggttctcagttgtgtgacaagaacactgtgtattatcagag 2218  
|||||  
Db 2184 GGCATGACCCGCTTACTGAAGGTTTCTAGTTGTGACAAAGAACACTGGTGTATTTACGAG 2243  
QY 2219 gacagttatgaagatatttcaagcaactgtgtgtaaaaaaaatgcatattgaacaaaga 2278  
|||||  
Db 2244 GACAGTTATGAGATATTTGACATACTTCTGTAGTATAAAAACAATGSCATTGAAACCAAGA 2303  
QY 2279 agcttctccagaatltcaagacacacttagcaatagcnaaaagcaattlaatgtccaccca 2338  
|||||  
Db 2304 AGCTTCCCAAGATTCOAAGACACCTTAGCATAGCAACGAAAGCAATTTATGCCCACCCA 2363  
QY 2339 ccagttcttgaagcccatcaacggaataactcgtactactcttcaagatcaagagag 2388  
|||||  
Db 2364 CCAAGTCTTGAAGCCCTTCAACGGGAATTAACCTGCTCTCTCTGATGAGTCAAGAG 2423  
QY 2399 gaaattgacatgtatgtatcacataltcagttgaaatgaagaagagaaatttgacattat 2458  
|||||  
Db 2424 GAATTTGACTATGATGATACCATATTCAGTTGAAATGAAGAAAGATTTTGACATTTAT 2483  
QY 2459 gatgaagatgaanaatcagaagcccccagacttltcaaaaagaanaacagacacattat 2518  
|||||  
Db 2484 GATGAGATGAATAATCAGAGCCCCCGACCTTCAAAAAGAAACACAGACTATTTTATT 2543  
QY 2519 gctgcagttgagagagcctctggtatltatgtggaatgagtgcccccacatgttctaaagaac 2578  
|||||  
Db 2544 GCTGCACTGGAGAGGCTCTGGGATATGGGATGATGAGTGCCTCCACATGTTCTTAAGAAAC 2603  
QY 2579 agggctcagagttgcaagtgctccctcagttcaagaagaattgttctcaggaattactgat 2638  
|||||  
Db 2604 AGGGCTCAGAGTGGCACTGCTCCAGTTCAGAGAAATGTGTTTCCAGGAATTTACTGAT 2663  
QY 2639 ggtctccttactcagcccttataccgttgaagaactaaatgaacatttgggaactcctggg 2698  
|||||  
Db 2664 GGCCTCTTACTCAGCCCTTATACCGTGAGAACTAATGAATGAACATTTGGGACTCCTGGGG 2723  
QY 2699 ccatataatgaagcagaagtttgaagataataatcatatgttacttccagaatcagcctct 2758  
|||||  
Db 2724 CCATATATTAAGAGAGAAAGTTGAAGATATATCTATGTTACTTTCAAAAACAGGCTCT 2783  
QY 2759 cgtcccatctcctctatacttcaagcctatltctataatgaagaatcgaagcagaagaca 2818  
|||||  
Db 2784 CGTCCCTATTCCTTCTATTTCTAGCCTTATTTCTTATGAGAAAGTCAAGAGCAAGAGCA 2843  
QY 2819 gaacctgaanaaaacttlttcaagcctaaltgaacccaanaacttlttgaagatgcaa 2878  
|||||  
Db 2844 GAACCTTGAAAAAACTTTGTCAAGCCTAATGAACCAAAAACCTTTTGGAAGATGCAA 2903  
QY 2879 catcataatgcaacccactaaagatgagttgtactgtcaaaacgctgggtctatctctcgtat 2938  
|||||  
Db 2904 CATCATATGTGCACCCACTTAAGATGATTTGACTGCAAAACCTGGGCTTATTTCTCTGAT 2963  
QY 2939 gtctgaacctggaanaagatgtgcactcagcgctgtatgtgaaccccttctgtctgcacact 2998

|||||  
Db 2964 GTTGACCTGGAAAAAGTGTCTCACTCAGCCCTGATGGACCCCTTGTGCTGCCACACT 3023  
OY aacacactgaacccctgtcctcattgaggagaacagtgacagtaacaggaattgctgtcttc 3058  
OY 2999 aacacactgaacccctgtcctcattgaggagaacagtgacagtaacaggaattgctgtcttc 3058  
Db 3024 AACCACTGAACCTGTCTCATGGAGACAGTACAGTACAGGAATTTGCTGTGTTTTC 3083  
OY 3059 aacacactgaacccctgtcctcattgaggagaacagtgacagtaacaggaattgctgtc 3118  
Db 3084 ACCATCTTTGATGAGACAAAAGCTGTACTTCACTGAAAATATGGAAGAAGATCGACGG 3143  
OY 3119 gctccctgcaatataccagatggaagatcccaatttaagagaattatccgtctcatgca 3178  
Db 3144 GCTCCGCAATATCCAGATGGAAGATCCACTTTTAAAGAAATATTCGCTTCATGCA 3203  
OY 3179 atcaatgcaataatgataacatacactgctgctgattgagtcagagatcaagaagatt 3238  
Db 3204 ATCATATGGCTACATTAATGATACACTACTGCTTAAATAGCTCAAGATCAAGATTT 3263  
OY 3239 cgaatgctatcctcagcatgagcgaatgaacatccatctcattcattcattcagtgga 3298  
Db 3264 CGATGATATCTGCTCAGATGGGCAATGAAAACATTCATTTTCACTTTCAGTGA 3323  
OY 3299 catggtctcactgtaagaaaaaagaggatataaatagcaatgcaatctcatcca 3358  
Db 3324 CATGTCTTCACTTACGAAAAAAGAGCATATAAATGGCAGCTGTACATCTCTATGCA 3383  
OY 3359 ggtgtctttagaagatggaatgtttacatccaaagcttggaatttgccgggtggaatgc 3418  
Db 3384 GGTGTCTTGGAGACAGTGAATGTTACCATCCAAAGCTGGATTTGGGGGTGGAATGC 3443  
OY 3419 ctatctgagagcaatcactatcactgctggaatgagcacaactttctggtgtaacagcataag 3478  
Db 3444 CTATATGGCGAGCATATCAATGCTGGGATGAGCAACATTTTCTGCTGATACGCAATAG 3503  
OY 3479 tctcagaactccctcgggaatggtctcggagacatagagatttcagattacagcttca 3538  
Db 3504 TGTGAGACTCCCTCGGGAGATGGCTTGGACATTAAGATTTTCAATTAACACTTCA 3563  
OY 3539 ggaacatgagagatgagcccaaaagcttgccagactctatctccggatccaatcaat 3598  
Db 3564 GGACAAATATGACAGTGGGCCCAAGAGCTGGCCAGACTTCAATTTCCGGATCAATCAAT 3623  
OY 3599 ggcctgagacccaaagaccccttctctggaatcaagtgatctgctgagccaaatgatt 3658  
Db 3624 GCCGAGAGACCAAGAGGCCCTTTCTTGATCAAGGTGATCTGTGGCAACCAATGAT 3683  
OY 3659 attcagggcatcaagaccagagtgcccgctcaagaattctccagctcttaactctcaag 3718  
Db 3684 ATTCAACGGCATCAAGACCCAGGCTGCCGTAGAAATCTCCAGCCTCTACATCTCTAG 3743  
OY 3719 ttatcatcatgatagtctctgagtggaagagatgagcagactctatccgaggaattccact 3778  
Db 3744 TTTATCTATCATGTATAGTCTTGTGATGGAGAAAGTGGCAGACTTATCGAGAAATTCCT 3803  
OY 3779 ggaacacttaatgctctccttggaatgagatggaatcattcggatataaaacacataattt 3838  
Db 3804 GGAACCTTAATGTCTCTTGTGCAATGTGATTCATCTGGATTAACACAAATATTTT 3863  
OY 3839 aacccctcaattatgctcgatataccgttttgcaaccaactcatatagcatccgagc 3898  
Db 3864 AACCTTCAATTAATGTCTGATCATCGTTTGCAACCACTTAATTAAGATTCGACAGC 3923  
OY 3899 actcttgagatgagtgatgagtgatgatttaaatagtttcagacatgcaattggaatg 3958  
Db 3924 ACTCTTTCATGATGAGTGTGATGGCTGTGATTTAATATGTTCAAGATTCATTTGGGAATG 3983  
OY 3959 gagaagtaaacataatagatgacagatctactgctcactcacttacaataatgctt 4018  
Db 3984 GAGAGTAAGCAATATACAGATGACAGATTAATCTGCTTACTTACTTACCAATATGTTT 4043  
OY 4019 gccacctggtctccctcaaaagctcgacctcaactccaaggaggagtaatgctcgagga 4078  
|||||

Db 4044 GCCACCTGCTCTCTTCAAAAAGCTCGACTTCACTCCAGGAGGAGTAATGCCGTGAGA 4103  
OY 4079 cctcaggtgaataatccaaaagatggtcgaagtggactccagaagaatgaagtgc 4138  
Db 4104 CCTCAGGTGAATTAATCCAAAAGAGTGGCTGCAAGTGGACTTCCAGAAACATTAAGTGC 4163  
OY 4139 acaaggaataactactaagggagtaaaatcctgcttaccagcatgatatggaagagctc 4198  
Db 4164 ACAGAGTAATCACTCAAGGAGTAAATCTGCTTACACAGCATGATGGAAGAGTTTC 4223  
OY 4199 ctcatctcagcagatgcaatgagatgagatgagatctcttttaagaatggaagta 4258  
Db 4224 CTCATCTCAGACGTAAGATGAGCCATCATGAGACTCTCTTTTTCAGAAATGGCAAGTA 4283  
OY 4259 aaggttttcagggaaacaaagactcctcaacactggtggaactctcagaccacag 4318  
Db 4284 AAGGTTTTTCAGGAGAAATCAACAGACTCTTCAACACTGTGGTGAACCTCTAGAACCCAG 4343  
OY 4319 ttactgactcgtacactcctgaattccaccagaagttggtgcaacagatgcccagag 4378  
Db 4344 TTACTGACTCGCTACTTCAATTCACCCAGAGTTGGTGACACAGATGGCCCTGAGG 4403  
OY 4379 atggaagttcgtggtcgtcgagagcaagacactctactgga 4417  
Db 4404 ATGAGGTTCTGGGCTGGAGGACACAGGACCTCTACTGA 4442  
  
RESULT 3  
108644 LOCUS 108644 4278 bp PAT 02-DEC-1994  
DEFINITION Sequence 4 from Patent WO 8800831.  
ACCESSION 108644  
VERSION 108644.1 GI:588650  
KEYWORDS  
SOURCE  
ORGANISM  
Unknown.  
REFERENCE  
1 (bases 1 to 4278)  
AUTHORS  
Pasek M.P.  
TITLE  
DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR  
VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE  
POLYPEPTIDES IN HIGH YIELDS  
JOURNAL  
Patent: WO 8800831-A 4 11-FEB-1988;  
FEATURES  
source  
1..4278  
location/Qualifiers  
BASE COUNT 1244 a 942 c 946 g 1146 t  
ORIGIN  
  
Query Match 93.9%; Score 4148.8; DB 10; Length 4278;  
Best Local Similarity 97.9%; Pred. No. 0;  
Matches 4270; Conservative 0; Mismatches 2; Indels 90; Gaps 3;  
  
OY 59 gccacacagaagatactactacctggtgagtgagtggaactgtcatgagacataatgcaagtgat 118  
Db 1 GCCACCAAGAAATACTACTGCTGGTGAAGTGAACGTGCATGAGGACTATATGCAAAAGTAT 60  
OY 119 ctggtgagcgtcgtcgtgagcgaagattctctctagagtgccaataatttctcatctc 178  
Db 61 CTCGGTGAAGTCCCTGTTGAGAGCAAGATTTCTCTGAGAGTCCAAATCTTTTCCATTC 120  
OY 179 aacacctcagtcgtgtacaaaagactctgttttagaatcaaggttcaaccttttaac 228  
Db 121 AACACCTAGTCTGTACAAAAGACCTCTGTTTGAATAATTCACGAGATCACTTTTTCAC 180  
OY 229 atcgctaaagccaaagcccttgatggtgctgtctgtaggtctcaaccatcagcagctgagtt 298  
Db 181 ATGCGTAAGCCAAAGCCACCTGGATGGGTCTGTAGGCTCTACCAATCCAGCGGTGAGGTT 240  
OY 299 tatgatagatggtcatatacacttaagaacatggtctccatccatctgctcaatgct 358  
Db 241 TATGATACAGTGTCTTATACATTAAAGACATGGCTTCCATCCTGTCATCTTCAATGCT 300  
|||||

OY 359 gttagtataccttacttgaaagcttctgagggagctgaataatgatatgacagacgtcaa 418  
|||||  
Db 301 GTTGTGTATCTCTAGTGAAGAGCTTCTGAGGAGCTGATATGATGATCGACACGACAA 360  
OY 419 aggagaagaagaatgaataagcttcctcctgctgtagaagccaatacatatgctgcagctc 478  
|||||  
Db 361 AGGAGAAAGAGATGATTAAGTCTTCCCTGCTGGAAGCCATACATATGCTGCGAGCTC 420  
OY 479 ctgaagaagaatggttccaatgaccttgagccactgtgaccttacctataatcttctc 538  
|||||  
Db 421 CTGAAAGAGATGCTCCAAATGCGCTCTGACCCACTGTGCTTACCTACTCATATCTTCT 480  
OY 539 catgtgagccgtgtaaaagacttgaatctcagccctcatgttagagccctcagtagataga 598  
|||||  
Db 481 CATGTGAGCTCGTAAAGACTTAATTCAGGCTCATTTGAGCCCTACTAGATATGTAGA 540  
OY 599 gaaggagctctgcgaagaagaacacagaccttgacaaatltatactacttctgct 658  
|||||  
Db 541 GAAGGAGTCTGGCCAAAGCAACACACACCTTGCACAAATTTATACTACTTTTGTCT 600  
OY 659 gtatttgtagaagggaagatgtgacctcagaaacaaagaacttctgtatgcagagatag 718  
|||||  
Db 601 GTATTGTATGAAGGAAAGTTGGCACTCAGAAACAAAGAACTCTTGATGACAGATAGG 660  
OY 719 gatgtcatcatgctgtgagcctgagcctaaatgtcacagatcaatggtatgttaacagg 778  
|||||  
Db 661 GATCTCATCTCTGCTCGGCTGGCTTAAATGCAACAGCAATGATTTATGTAAACAGG 720  
OY 779 tctctg---cgaagctcgtatgtgtagcacaggaatcagctctatgtgcatgtatgga 835  
|||||  
Db 721 TCTGTGTACCAAGTCTGATTTGATGTCACAGAAATCAGTCTTTGGCATGTGATTTGGA 780  
OY 836 atggagcaactcctcgaagtgactcaatatctcctgaaggtccacacatttctgtgag 895  
|||||  
Db 781 ATGGGCACTCTCTGAAAGTCACTCAATATTCCTCGAAGTCAACATTTCTGTAGG 840  
OY 896 aacatctgcgaaggtctcttggaatcttcgcgaataacttcttactgtctaaacactc 955  
|||||  
Db 841 AACCATGCGCAGGCTCTTGGAATCTCGCANTATCTTCTACTGCTCAACACTC 900  
OY 956 ttgatgaccttgagacagtttctactglttctgcatatactcttcccaacaatgatgagc 1015  
|||||  
Db 901 TTGATGACCTTGAGAGTTCATCTGTTTGTGATATCTCTTCCACCAACATGATGGC 960  
OY 1016 atggaagcttatgtcaaatgtagacagctgtccagaggaaccccaactagaatgaaaaat 1075  
|||||  
Db 961 ATGGAACCTTATGTCAAAAGTAGACAGCTGTCCAGAGGAACCCCAACTACGATGAAAAAT 1020  
OY 1076 aatgaagaacggaagaactatgatgatcttactgtatcttgaatgtgagtgtgtag 1135  
|||||  
Db 1021 AATGAAAGAGCGGAAGACTTATGATGATGATCTTACTATTTCTGAAATGTAGTGTAGG 1080  
OY 1136 ttgatgatgacaactcctccttcttataccaatctgcctcagttgcgaagaagcactcct 1195  
|||||  
Db 1081 TTTGATGATGACAACCTCTCTCTTATTCOAATTCGTCAGTTGCGCAAGAACATCTCT 1140  
OY 1196 aaaacttgggtacattacattgtctgtcgaagaggaagacttggagactatgctccctagtc 1255  
|||||  
Db 1141 AAAAATTGGGTACTATTCATTTGCTGTGAAGAGAGAGACTGGGACTATGCTCTCTAGTC 1200  
OY 1256 ctgcgccccgatgacagaagtltataaagtccaatlttgaaacatgacctcagcgagat 1315  
|||||  
Db 1201 CTCGCCCCGATGACAAAGTATTAAGTCAATATTGAAACAATGGCCCTCACGGATTT 1260  
OY 1316 gtagagaagtacaaaaaagtcocgaatttatgtcatacacagatgaacacttlaagacct 1375  
|||||  
Db 1261 GGTGGAAGTACAAAAAAGTCGGATTTATGCGATACACAGATGAACCTTTAAGACCTGT 1320  
OY 1376 gaagctatcaagcatgaatcaaggaactcttgggaccttacttcatggggaagtttgagac 1435  
|||||  
Db 1321 GAACTATTTCAGCATGAATCATGAGAACTTTGGACCTTACTTTATGGGGAAGTTGGAGAC 1380

OY 1436 aacatgtgatatattatattaaagaatcaagcaagcagaccataatacatctacccctcagga 1495  
|||||  
Db 1381 ACAGCTTTGATTTATTTAAAGATCAAGACAGACCATATACATCTACCTCACGGA 1440  
OY 1496 atcaatgatlccgltcccttctglatccaaggagatctaacaaaggltgtaaacatttgaag 1555  
|||||  
Db 1441 ATCAGTATGTCGCTCTTTGTATTTCAAGGAGATTACCAAAAGGTGTAAMAACATTTGAAG 1500  
OY 1556 gatttccaattctgcagagagaataatccaataataatgagcagtgcactgtagaagat 1615  
|||||  
Db 1501 GATTTTCCAAATTCGCCAGAGAAATATTCAAAATTTAAATGGACAGCTGTAGAAGAT 1560  
OY 1616 gggcccaactaaatcagaatccctcgtgtcctgaaccgccttaacttaacttgcgttaaatg 1675  
|||||  
Db 1561 GGGCCAACTAATATAGATTCCTCGGTGCTGACCCGCTATTACTTACTTTCGTTAATATG 1620  
OY 1676 gagaagatctagcttcaaggactcatgtgaccttccctcatctgtctacaagaatcgtga 1735  
|||||  
Db 1621 GAGAGAGATCTAGCTTCAGGACCTCATTTGGCCCTCTCCATCTGCTCAAAAGAAATCTGTA 1680  
OY 1736 gatcaaaagaggaacacagataatgtcaagaagaagaaatgcatcctgttctcgtatct 1795  
|||||  
Db 1681 GATCAAAAGAGAAACCAAGATATGTCAAGAACAGGAATGTCATCCTGTTTCTGTATTT 1740  
OY 1796 gatgaagaccgaagctgtgacctcacaagagaataatacaagcttctcccaatccagct 1855  
|||||  
Db 1741 GATGAGAACCGAAGCTGTGACTTCTACAGAGATATACAGGCTTTCCTCCAAATCCAGCT 1800  
OY 1856 ggaagtcagcttggaga tccaagatltccaagctcacaacatcatgacaagcatcaatgagc 1915  
|||||  
Db 1801 GAGATGAGCTGTGAGATCCAGATTCAGAAAGCTCCCAACATCATGACAGATCAATGGC 1860  
OY 1916 tatgttlttataagtttgcagttgtcagttgttcttgcatagagttgcatctgttactc 1975  
|||||  
Db 1861 TATCTTTTGTATGATTTGCAATGTCTCAGTTGTGTTGATGATGAGGCGATCTGTACTATT 1920  
OY 1976 ctgaagcatlgagacagagactgacttctctctgtcttcttctcttgatibactticaa 2035  
|||||  
Db 1921 CTAAAGCTTTGGAGCAGACTGACTTCTTCTTCTTCTTCTTCTGATATACCTTTAAA 1980  
OY 2036 cacaaaaatgttctataagaacacacactacccatltcccatctcaagagaacatgtctc 2095  
|||||  
Db 1981 CACAAATGCTCTATGAAAGACACACTCACCTATTCCATTTCTCAGAGAAACTGTCTTC 2040  
OY 2096 atgtctgatlgaaaaacccaggttctatgtgattctggtgtgcacaactcaagacttccggaac 2155  
|||||  
Db 2041 ATGTGATGGA AAAACCCAGGCTATGTGATTCGTGGGTGCCAACACTCAGACTTTCGGAAC 2100  
OY 2156 aagaagcatgaccgccttactcgaaggttctcagttgtgacaagaacacttgtgatatctac 2215  
|||||  
Db 2101 AGAGGCTAGACCGCTTACTGAAAGTTTCTAGTTGTGACAAAGAACACTGGTGATTTATTC 2160  
OY 2216 gaagacagttatgaagatatlttcagcatctgtctgtagtaaaaaaaatgtccattgaacca 2275  
|||||  
Db 2161 GAGGACAGTTATGAAGATATTTCAGCATCTGCTGTGATTA AAAACATATGCGATGAAACA 2220  
OY 2276 agaaagcttctcccaagaatltcaagaacccctagcaactagcgcaaaagcaatttaagccacc 2335  
|||||  
Db 2221 AGA----- 2223  
OY 2336 ccaccagctcttgaagcgcataccaaggaataactcgtactactcttcagtcagatcaa 2395  
|||||  
Db 2224 -----GAATATCTGTCTACTCTTTCAGTCAAGATCAA 2256  
OY 2396 gaagaaatlgactatgtaga laccata tcaatgtgaaatgaagaagagaatlttgacat 2455  
|||||  
Db 2257 GAGGAATTTGACTATGATGATACATATCAGTTGAATGGAAGAAAGATTTTGAACATT 2316  
OY 2456 tatgatgagtagtgaataatcaagagccccgcagcttltcaaaaaagaacagacactatltt 2515  
|||||  
Db 2317 TATGATGAGATGAAATTCAGAGCCCGCAGCTTTTCAAAAAGAAACACACACTATTTTT 2376  
OY 2516 atgctgcagtgagagagcctctgagatctatgtagatgtagctccccaatgttctcaga 2575



|||||  
Db 2377 ATTCTGAGTGGAGAGCCTCGGGATTATGGATGAGTAGCTCCCAATCTTTAANA 2436  
2576 aacgggctcagagtgagcagtgctccagtgctcaagaagaagtgcttccaggaatttact 2635  
|||||  
Db 2437 AACGAGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTTCCAGAAATTTACT 2496  
2636 gatgctccttactcagcccttatacgttggagaactaaatgaaacatttggagactctg 2695  
2497 GATGGCTCCTTACTCAGCCCTTATACGTGGAATAATGAACATTGTGGAGCTCG 2556  
Qy 2696 gggccataataagaagcagaagttgaaataataatcatgtaacttcaagaataagcc 2755  
2557 GGGGCATATATAGAGCAGAAAGTTGAAATATATCATGTAACTTTCAAGAAATCAGGCC 2616  
Qy 2756 tctgctccttactccttacttacttacttacttacttacttacttacttacttacttact 2815  
2617 TCTGCTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2676  
Qy 2816 gcgaacctagaaaaaacttctcagagcctaatagaaccaaaacttacttctgaaagtg 2875  
2677 GCAGACCTAGAAAAAATTTGTCAAGCCTATGAACCAAACTTACTTTTGGAAAGTG 2736  
Db 2876 caacataatctgcaaccacataaagaatgagttgacgtcaaaagcctgggttacttctc 2935  
2737 CAACATATATGACACCCACCTAAAGATGAGTTGACTGCAAAAGCCTGGCTTATTTCT 2796  
Qy 2936 gatctgacccggaaaaaagatctgacacagccctgacttggagcccttctgtctgcaac 2995  
2797 GATTTGACCTGGAAAAAGATGTGCACCTCAGGCCCTGATTTGGACCCCTTGTGCTGCCAC 2856  
Qy 2996 actaacaactgaacctctcacttggagagaagtgacagtaacgaatctgtctgtt 3055  
2857 ACTAACACACTGAAACCCCTGCTCATGTGGAGACAAAGTAGACAGTAAGAAATTTCTGT 2916  
Qy 3056 t---tcacacacttctgtagaagcaaaagctggttacttccagaaataatggaagaac 3112  
2917 TTCTCAGCATTCTTTGATGAGAACCAAAAGCTGTACTTCACTGAAATAATGGAAGAAAC 2976  
Qy 3113 tgcagagctccctcgaatatccagatggaatgccactttaaagagaatctgctc 3172  
2977 TGCAGGCTCCTCCTCAATATCCAGATGGAAGATCCCACTTTTAAAGAGAAATTTATGCTTC 3036  
Qy 3173 catgcaatcaatggctacataatgatacactggtctgtaatggtctcagaatcaa 3232  
3037 CATCAATCAATGGCTACATATGATGATACACTGCTGCTTATGATGGCTCAGATCAA 3096  
Qy 3233 aggatctgagtgatctctcagcaatggcagcaatgaaataatccatcttcttcaattc 3292  
3097 AGGATTCGATGGTATCTGCTCAGCATGGGCAAGCAATGAACATCCATTCTATTTCTTC 3156  
Qy 3293 agtgagacatgctcactgtacgaaaaaaagagagtaaaatgacactgtacacatcc 3352  
3157 AGTGCACATGTGTCTCAGCTGACGAAAAAAGAGAGATTAATAATGCGCATGTACATCTC 3216  
Qy 3353 taccagggtgttcttgagacagtgaaatgltacacatccaaagctggaatttgccggtg 3412  
3217 TATCCAGAGTGTTTTGGAGACAGTGAATGTATCCATCCAAAGCTGGAATTTGGGGGTG 3276  
Db 3413 gaatgcttatttggcagcatctacatgctgagatgacaaacttcttctgtgtaagc 3472  
3277 GAATGCCCTTATTTGGCAGCATCTACATCTGGGATGACACACTTTTCTGTGTACAGC 3336  
Qy 3473 aataagtgacagactcccttggaaatgagctcttgagacatataagatcttcaattaca 3532  
3337 AATAAGTGTGAGACTCCCTTGGGAATGGCTTCTGACACATTAAGATTTTCAGATTACA 3396  
Qy 3533 gcttcagagacaataatgagcagtgggcccaaaagctgcccagactcatatccgagta 3592  
3397 GCTTCAGAGCAATATGACACTGGGCCCAAAAGCTGCCAGACTTCATTATTCGGATCA 3456  
Qy 3593 atcaatgcttgagacacaaagagcccttcttcttgatcaagtgatctgttgacaca 3652  
|||||

Db 3457 ATCAATGCCCTGGAGACCAAGAGGCCCTTTCTTGGATCAAGTGATCTGTGGCACCA 3516  
Qy 3653 atgatattcaagcatcaagaaccagggtgcccgtcagaagttctccagctctacatc 3712  
|||||  
Db 3517 ATGATTATTCAGGCATCAAGACCCAGGGTCCCGTCAAGAGTTCTCAGGCTTATCATC 3576  
Qy 3713 tctcagttatcaatcaglatagcttgaatggagaagtgccagacttalcagaagaat 3772  
3577 TCTCAGTTATCATCATGTATAGTCTTGATGGAGAAAGTAGTGCGACATTATCGAGAAAT 3636  
Db 3773 tccactggaaaccttaatgcttcttcttggcaatgtaataatcttggatataaacaacat 3832  
3637 TCCACTGGAACCTTATATGTCTTCTTGGCAATGTGATTCATCTGGATTAACACCAAT 3696  
Qy 3833 attttaaccctccaatatttgcctgatacaccgttgcaccccaactatataagcat 3892  
3697 ATTTTAAACCTCCCAATTTATGCTGATACATCCGTTTGGACCAACTCATATAGCAT 3756  
Db 3757 CCGACACTCTTCGATGAGTGTGATGGCTGTGATTAATAGTTGACAGATCCATTG 3816  
Qy 3953 ggaatggagatgaagaataatcagatgcagagttactgttacttacttacttact 4012  
3817 GGAATGGAGAGTAAAGCAATATCATATGCACAGATTACTGCTTATCTTACTTTACCAAT 3876  
Qy 4013 atgttgcacactgttcccttcaaaagctgacttcaaccctcaaggaaggaatgagc 4072  
3877 ATGTTGGCACCTGTGTCTCTTCAAAAGCTGACTTCACTCCCAAGGAGAGATGATGCC 3936  
Qy 4073 tggagaacctcagtgataatccaaagaatggtcgtcaagtgtgacttccagaagaatg 4132  
3937 TGGAGACCTCAGGTGAATAATCAAAAGAGTGGCTGCAAGGAGACTTCCAGAAACAAATG 3996  
Qy 4133 aaagtccacagaatgtaactcactcagggagtaaaactctgcttcaacagatgtagaag 4192  
3997 AAAGTCACAGAGTACTCTCAGGAGTAAATCTGTGCTTACACAGATGATGTGAAG 4056  
Qy 4193 gagttccctacatccagcagctcaagatgagcatcagtgagactcttcttccagaatg 4252  
4057 GAGTTCCTCATCTCCAGAGTCAAGATGCCATGAGAGACTCTTTTTCAGAAATGCC 4116  
Qy 4253 aaagtaaaagtttctcagggaaatcaagaactctcactcagctgtgtgtaactctagac 4312  
4117 AAAGTAAAGGTTTTTCAAGGAAATCAAGACCTCTTCAACCTGTGGTGAACCTCTAGAC 4176  
Db 4313 ccacgttactgactcgttactccttcgaattacccccagaagtggtggtcagcagatggc 4372  
4177 CCACGTTACTGACTCGCTTACTCTTGAATTCACCCCAAGAGTGGGTGACACAGATTGCC 4236  
Qy 4373 ctgaggaatgaggttctggtcgtcagagcagagactctac 4414  
4237 CTGAGGATGGAAGTCTGTGGCTGGAGGACACAGACCTCTAC 4278  
Db 4237 CTGAGGATGGAAGTCTGTGGCTGGAGGACACAGACCTCTAC 4278

RESULT 4  
108643  
LOCUS 108643 4281 bp  
DEFINITION Sequence 3 from Patent WO 8800831.  
ACCESSION 108643  
VERSION 108643.1 GI:588649  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4281)  
AUTHORS Pasek,M.P.  
TITLE DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE POLYPEPTIDES IN HIGH YIELDS  
JOURNAL Patent: WO 8800831-A 3 11-FEB-1988;  
FEATURES Location/Qualifiers  
Source 1..4281



BASE COUNT 1245 a 943 c 946 g 1147 t  
ORIGIN  
Query Match 93.9% Score 4147.2; DB 10; Length 4281;  
Best Local Similarity 97.9% Pred. No. 0;  
Matches 4269; Conservative 0; Mismatches 3; Indels 90; Gaps 3;  
QY 59 gccaccagaataataactactctggtggtgcagtggaactgtcatggaagctatatatgcaagaagat 118  
|||||  
DB 4 GCCACCGAAGAAATACTACTCTGGGTGCGAGTGGAACTGTCATGGGACTATATCCAAAGGAT 63  
QY 119 ctggttagctgctctggaagcaagaatttcctctagagtggaacaaatctttccattc 178  
|||||  
DB 64 CTCGGTAGCGCTCTGTGAGCGCAAGATTCTCTCTAGAGGCCAAATCTTTCCATTTC 123  
QY 179 aacacctcagtcgtgtacaaaagaactcgtttgtagaatcacggttcacctttcaac 238  
|||||  
DB 124 AACACCTCAGTCGTGTACAAAAGACTCTGTTGTAGATTCACGATCACCTTTTCAAC 183  
QY 239 atgcctaagccaagccaacctgtgagtgtctgtctaggtctaccaccagagcgaagtt 298  
|||||  
DB 184 ATGCGCTAAGCCCAAGGCCACCTGGATGGTCTGTAGGTCTCTACCATCCAGGCTGAGGTT 243  
QY 299 tatatacagtggtcatctacacttaagaacatggtctccactcctgtcaggtctcatgct 358  
|||||  
DB 244 TATGATACAGTGTCATTTACCTTAAGAACATGSGCTTCCCTCTGTGAGTCTTTCAATGCT 303  
QY 359 gtgggtatctcacttggaagctctctggaagctgtaataatga tga tcaagaccagtc 418  
|||||  
DB 304 GTTGGTGTATCTACTGGAAGGCTTCTGAGGAGCTGAATATGATGATCAGACCAAGTCAA 363  
QY 419 agggagaagaagaatgtataaagtctccctggtgtgaagccaataatatgtctgcaagtc 478  
|||||  
DB 364 AGGAGAAAGAAAGATGATTAAGTCTCCCTGGTGAAGCCATACATATGCTGTGCAAGTTC 423  
QY 479 ctgaaagaagaatgtgccatggtcctctgacccactgtgcttactctactatctctct 538  
|||||  
DB 424 CTGAAAGAGAAATGTCCTCAATGCGCTCTGACCCACTGTGCTTACTCTCATATCTTTCT 483  
QY 539 catgtgacctgtgtataaagaacttgaaatcagggcctcatctggaagccctactagta tga 598  
|||||  
DB 484 CATGTGAGACCTGCTTAAGAAAGCTTAATTCAGGCTCATTTGGAGCCCTACTAGTATGAGA 543  
QY 599 gaaggaggtctggccaagaaagaacacagaccttgcaaaaattatactactctttgct 658  
|||||  
DB 544 GAAAGGAGTCTGGCCAAAGGAAAGACACACACCTTTGCACAAAATTTATACCTTTTGTCT 603  
QY 659 gtaattgataaggaaagaatgtgacactcaagaacaaagaactccttgatgcaagatag 718  
|||||  
DB 604 GTATTGTATGAAGGAAAGATGTGGCACTCAGAAACAAAGAACTCTTGATGACGAGATAG 663  
QY 719 gatgctgcatctgctgggctggtcctaaatgtcacacagtcacatggtatgtatgaacag 778  
|||||  
DB 664 GATGCTCATCTGCTCGGGCTGGCTTAATAATGCACACAGCATATGTTATGTAAACAGG 723  
QY 779 tctctg----ccaggtctgattgtgatagtccaaaggaatcagtctctatgtgcatgtgatgga 835  
|||||  
DB 724 TCTGTGCTACCAAGGTCTGATTGTGATGTCACAGGAATCAGCTATTTGGCAGTGTATTGGA 783  
QY 836 atgggacacactcctgaagtgcaactcaatatctcctgaagtgcaacatctctgttagg 895  
|||||  
DB 784 ATGGGACACACTCTCTGAAGTGCACTCAATATTCCTCGAAGTGCACATTTCTTGTAAG 843  
QY 896 aaacatcgcaagcgctctctgtgaaatctgcaccaataactctcctactgtcacaacatc 955  
|||||  
DB 844 AACCATGCGCAGGCGCTCTTGGAATCTCGCATTAATCTTCTTACTGTGCAAAACATC 903  
QY 956 ttgtagaccttgacaagtttctactgttttgtcatatctctctccacaacatgatagggc 1015  
|||||  
DB 904 TTGTATGAGACCTTGAGAGTTTCTACTGTGTTTGTATCTCTTCCACCAACATGATGGC 963

QY 1016 atggaagcttatgtcacaagtacagcgtgtccaggaagaccaccaactacagatgaataat 1075  
|||||  
DB 964 ATGGAAGCTTATGTCAAGATGACAGCTGTCCAGAGAACCCCACTACGATCAAAAAAT 1023  
QY 1076 aatgaagaagcgggaagacatgatgatcttactgactctgaaatgtatgtgttcaagg 1135  
|||||  
DB 1024 AATGAAGAAGGGAAGACTATGATGATGATGATCTTACTGATCTGAAATGATGTGTGACG 1083  
QY 1136 ttgatgatgaacaactctcctctctctatccaatttgcctcagttgtgcaagaagcatcct 1195  
|||||  
DB 1084 TTGTATGATGACAACTCTCTCTCTTATCCAAATTTGCTCAGTTGCTCAAGAAACATCTCT 1143  
QY 1196 aaacttgggtacatacatgtcgtcgtgaagaagagctggtgactgtcctcctatgtc 1255  
|||||  
DB 1144 AAACTTGCGGTACATTATGCTGTGAAAGAGAGACTGTGGACTATGCTCCTTATGTC 1203  
QY 1256 ctgcgcccgatgaagaaagtataaagtlcaatatttgaacaaatggccctcagcgatt 1315  
|||||  
DB 1204 CTCGCCCCCGATGACAGAAAGTTATAAAGTCAATATTTGAACAATGGCCCTCAGCGATT 1263  
QY 1316 ggttagaagtaacaaaagtcggattatgtgcatcacagatgaacaccttaagaactcgt 1375  
|||||  
DB 1264 GGTAGGAAGTACAAAAGTCCGATTTATGGCATACAGATGAACCTTTAAGACTCGT 1323  
QY 1376 gaagctatcaagcatgaatcaggaatcttggaacttacttlatggtgaagttgagac 1425  
|||||  
DB 1324 GAACCTATTGACATGATGATACGAATCTTGGAACCTTACTTTATGGGAAAGTTGGAGAC 1383  
QY 1436 aacacttgatlatatattaaagaatcaagaagcagacataataacatcactcctcagga 1495  
|||||  
DB 1384 ACACTGTTGATTATATTAAAGATCAAGCAAGCAGACATTAACATCTACCTCCACGGA 1443  
QY 1496 atccatgattcgcgtctcttggatctcaagagatctaccaaaagtgataaacaatttgaag 1555  
|||||  
DB 1444 ATCACTGATGTCGCTCTTGTATTAAGGAGATTACCAAAAGGTGTAAACATTTGAAG 1503  
QY 1556 gatttccaattctgcagcaggaataatccaatataatggaagtgatgcaagat 1615  
|||||  
DB 1504 GATTTTCCAAATCTGCCAGAGAAATATTCAAAATATAATGACAGTGACTGTGAAGAT 1563  
QY 1616 gggccaactaaatcaagatcctcgtgtcctgacccgctactactagtttcgttaatatg 1675  
|||||  
DB 1564 GGGCAACTAAATCAGATGATCCGCGTGCCTGACCCGATATTCTGATTTGCTTAATATG 1623  
QY 1676 gaaggaatctagcttgaaggaactatggccctctctctactctgctcaagaagatctgta 1735  
|||||  
DB 1624 GAGAGATGCTAGCTTCAAGGACTATTGGCCCTCTCTCATCTGTCTCAAAAGAAATCTGTA 1683  
QY 1736 gatcaagaagaacacagataatgttcaagaagaagaatgtcatcctgcttctgtaatt 1795  
|||||  
DB 1684 GATCAAAAGAGAAACCAAGATATGTACAGACAGAGATGTCAATCCGTGTTTGTGATTTT 1743  
QY 1796 gatvgaaccggaagcgtgtacactcaagagaataatacaagccttctcccaatccagct 1855  
|||||  
DB 1744 GATGAGAACCGGAAGCTGTACTCTCAGAGAAATATACAAAGCTTTTCCCAATCCAGCT 1803  
QY 1856 ggaatgaagcttgaagatgtccaaaggttccaaagccccaagctcaatgaagcaatcaatggc 1915  
|||||  
DB 1804 GGAAGTCAAGCTTGAGATGCCAGATTCGAAGCTTCAAACTCAATGACACAGCATTAATGGC 1863  
QY 1916 tatgttttatagttgtcagttgttctgataagtgatgagtgatcactgtaact 1975  
|||||  
DB 1864 TATGTTTGTATAGTTTGCAGTTGTGATGTTTGTGATGATGAGTGCGATACTGGTATATT 1923  
QY 1976 ctaagcatgtgagcaagactgactcctctctgtctctctctctctctctctctcaaa 2035  
|||||  
DB 1924 CTAAGCATTTGGAGACAGACTGACTTCTTCTGTCTTCTCTCTGATATACCTTCAAA 1983  
QY 2036 caaataatgtctatgaagcaacatcaacatcttccactctcaaggaanaacgtcttc 2095  
|||||  
DB 1984 CACAAATTTGCTATGTAAGACACATCACCTATTTCCCATTTCTCAGGAGAACTGTCTTC 2043  
QY 2096 atgtcgaatgaaaccaggtctatgtatctctgtggtgtgcacaacactcagcttctggaac 2155

|    |      |  |      |
|----|------|--|------|
| Db | 2044 | ATTTGCATGGAAAAACCAAGGCTCTATGATTTCTGGGTGCCCACTACAGACTTTGGGAAC       | 2103 |
| Qy | 2156 | agaagcaltgacccgcttacttgaagaagtttctagttgtgacaaagaacctgltatcatc      | 2215 |
| Db | 2104 | AGAGGATGACCGGCTTACTGAAAGGTTTCTAGTTGTGACAAAGAACACTGGTGATTTATAC      | 2163 |
| Qy | 2216 | gagagacgttaataagaatttcagcatactgtctgagtaaaacaatgacattgaacca         | 2275 |
| Db | 2164 | GAGGACAGTTATGAAAGTATTCTCAGCATACTCTGAGTAAATAAACAATTCATTCACCA        | 2223 |
| Qy | 2276 | agaagcttcctccagaattcaagacacccctagcactagcgaaagcaattaatgccaac        | 2335 |
| Db | 2224 | AGA-----   | 2226 |
| Qy | 2336 | ccacccagctctgaagcgcalcaacggyaanaactcgtactactctcagtcacatcaa         | 2395 |
| Db | 2227 | -----GAATTAAGTCTGCTACTACTCTTCAAGTCATATCAA                          | 2259 |
| Qy | 2396 | gaagaaattgaactatgataccatatcagttgaaatlgaaagaagaatttgaacatt          | 2455 |
| Db | 2260 | GAGGAATTTGACTATGATGATACCATATCAAGTTGAAATGAAGAAGATTTTGACATT          | 2319 |
| Qy | 2456 | tatgatatgagatagaataatcagagcccccgcagcttccaagaagaacacgacacattt       | 2515 |
| Db | 2320 | TATGATGAGGATGAAATCAGAGCCCGCAGCTTTCAAAAGAAACGACACTATT               | 2379 |
| Qy | 2516 | attgcttgcaatgagagccctctgagatataggaatgaatagctccccaatgtctaa          | 2575 |
| Db | 2380 | ATTGCTGCACTGGAGAGGCTCTGGGATTTATGGGATGAGTACTCCCACTGTTTAA            | 2439 |
| Qy | 2576 | aacagggctcagagtggtcagtgctccctcagltcaagaagaagttgtttccagaagattact    | 2635 |
| Db | 2440 | AACAGGGCTCAGAGTGCGACAGTCCCTCAGTTCAAGAAAGTTGTTTCCAGGAATTTACT        | 2499 |
| Qy | 2636 | gatggctcccttactcagcccttataccgttggaagaactaaatlgaaactttgagactctg     | 2695 |
| Db | 2500 | GATGGCTCTCTTCTCAGCCCTTATACCGTGAGAACTTAAATGAACAATTTGGGACTCTG        | 2559 |
| Qy | 2696 | gggccaataataagagagagaagtggaagataataatcagtgtaacttcagaaatcagcc       | 2755 |
| Db | 2560 | GGGCCATATTAAAGAGAGAGAAAGTTGAAGATTAATCATGATGTAACCTTTCAGAAATAGGCC    | 2619 |
| Qy | 2756 | tctcgtccctactctctctatcttctagccttatctctatagaagaatcagagcgaa          | 2815 |
| Db | 2620 | TCTCGTCCCTATTCCCTCTATTCTAGCCCTTATTCTTATGAGAAAGATCAAGAGCGCAAGA      | 2679 |
| Qy | 2816 | gaggaacctagaanaaaacttgttcaagcctaaatgaacccaacttactcttgaagaatg       | 2875 |
| Db | 2680 | GCAGGAACCTTAGAAAAAACTTTGTCAACGCTTATGAAACCAAAACTTACTTTTGGAAATG      | 2739 |
| Qy | 2876 | caacatcataatggcacccactaaagaatgagtttgacatgcaagccctgggtctatctct      | 2935 |
| Db | 2740 | CACATCATATATGGCACCCACTTAAGATGAGTTTACTGTCAAAAGCTTGCGCTTATTTCTCT     | 2799 |
| Qy | 2936 | gagtgtgaccttggaanaaagaatgtgacactcagcgcttgatltggaacccctctgltcgtcac  | 2995 |
| Db | 2800 | GATGTTGACTTGAAAAAGATGTGCTCACTCAGGCCCTGATTTGGAACCCCTCTGCTGCCAC      | 2859 |
| Qy | 2996 | actaacacactggaacccctgtctcatatggagagacagtgacagtaacagaaattgtctgttc   | 3055 |
| Db | 2860 | ACTTAACACACTGGAACCTCGCTCATGGGAGACAAAGTACAGTACAGGAATTTGCTCTGTTT     | 2919 |
| Qy | 3056 | t-----taacatcttttgaatgagacaaagaagctggtactcttacttgaanaataatggaagaac | 3112 |
| Db | 2920 | TTTCTCTACCACTTTTGTGTGAGAACCAAAAGCTGTGTAATTCCTACTGAAATAATGGAAGAAC   | 2979 |
| Qy | 3113 | tgcagggctcccccgaatatccagaatggaagatcccaactttaaagagaattatcgtctc      | 3172 |
| Db | 2980 | TGAGGGCTCTCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAAATTTATGCTTTC      | 3039 |
| Qy | 3173 | catgcaatacaatgggtacataaataggataccctactcctgtatgaatggtcccaagataca    | 3232 |
| Db | 3040 | CATGCATTCATAGGCTACTAATATGATACACTGCTGGCTTAGTAAAGGCTACAGATCA         | 3099 |
| Qy | 3233 | aggattcgaatglatcgtctcagcatggtggcagcaaatgaaacatccattctatcatc        | 3292 |
| Db | 3100 | AGGATTCGATGATCTGCTCAGCATGGGACACATGAAACATCCATTTATTCATTTTC           | 3159 |
| Qy | 3293 | agtggacatgtgttcaactgtaagaaaaaagggtgtataaatggaactgttacaatctc        | 3352 |
| Db | 3160 | ACTGGACATGTTTCTCATCTGCAAGAAAAAGAGGATTAATAATGCACTGTACATCTC          | 3219 |
| Qy | 3353 | tatccaggtgttttgaagaatggaagaatgttaccatcccaagaagctgtgaatttggcggtg    | 3412 |
| Db | 3220 | TATCCAGGTTTGTGAGACAGTGGAAATGTTACATCCAAAGCTGGAAATTTGGCGGGTG         | 3279 |
| Qy | 3413 | gaatgcttatltggtcgaacatcatatgcttggagatgaacacttltctgtgtacagc         | 3472 |
| Db | 3280 | GAATGCCATTATGGGAGAGCATGTACATGCTGGGATGAGACACTTTTCTGTGTACAGC         | 3339 |
| Qy | 3473 | aataagttcgaactccctctgggaatgtgtctcttgacaacatlagagaatttccaagataca    | 3532 |
| Db | 3340 | AATAAGTGCACAGCTCCCTGGGATGAGCTTCTGACACATTTAGATTTTACATTTACA          | 3399 |
| Qy | 3533 | gttcaagaagaatgatggaacatgtgggcccccaagctggcccaagactcatatccggaata     | 3592 |
| Db | 3400 | GCTTCAGCAATATGACAGAGGGGCCCAAGGCTGGCAGCTTCATTTATCCGAGTCA            | 3459 |
| Qy | 3593 | atcaatgctctgagacacaagaagagcccttctcttgatccaagtgatctgttgaccaca       | 3652 |
| Db | 3460 | ATCATATGCTGTGAGCACCAGAGAGCCCTTTTCTTGATCAAGTGTGATCTGTGGACCA         | 3519 |
| Qy | 3653 | atgattatcacgcatcaagaacagaggtgtccgctcagaagttctccagcctctacac         | 3712 |
| Db | 3520 | ATGATTTATTCACGGCATACAGCCAGGCTGCCCTCAGAAATTCACAGCTCTACATTC          | 3579 |
| Qy | 3713 | tctcagttatcatatcatatgacttctatgttggaagaaggttggaagactctcaggaagat     | 3772 |
| Db | 3580 | TCTCAGTTTATCATATGATATATGATTTGATGGAGAAAGATGGCAACTTTATCGAAGAAAT      | 3639 |
| Qy | 3773 | tccacttgaacattatgtcttctcttcttgccaatgtgatatcctcggagatataaacacat     | 3832 |
| Db | 3640 | TCCACTGGAACTTAAATGTTCTTTTGGCAATGTGATTCATCTGGGATTAACACAT            | 3699 |
| Qy | 3833 | attttaaaccctccaattatgtctcagatatacatccgtttgcaaccaatcatatagaact      | 3892 |
| Db | 3700 | ATTTTAAACCTCCCAATTATGCTCGATACATCCGTTTGCAACCCAACTTAATAGCATTT        | 3759 |
| Qy | 3893 | cgcagacactctgcagatgagttgagtggtgtgtgtttaaabaagtgtgagatccactg        | 3952 |
| Db | 3760 | CGCACACTCTTGCATGAGTGAATGGGCTGTGATTTAAATAGTTGACGATCGCATTG           | 3819 |
| Qy | 3953 | ggaatgagagatgaagaacatatacaatgtcaagatatctactgtctcatcttaccat         | 4012 |
| Db | 3820 | GGAATGGAGAGTAAGCAATATACATATGCACAGATTACTGCTCATCTTACTTTACCAT         | 3879 |
| Qy | 4013 | atgtttgccaacctgtctcctctcaaaagctcgaacttcaacctccaaggagagataatgccc    | 4072 |
| Db | 3880 | ATGTTTGCACACTGTGCTTCCTTCAAAAGCTGCATTCACCTCCAAAGGAGAGATATGCC        | 3939 |
| Qy | 4073 | tggagacctcaggtgaataataatccaaagaatgtgctgcaatgagacttccagaagaacatg    | 4132 |
| Db | 3940 | TGGAAGCTTCAGGTGAATTAATTCAAAAGAGTGGCTGCAAGTGGATTCACAAAGACATG        | 3999 |
| Qy | 4133 | aaagtacacagagtaacttcaaggaataaactctgcttaccagatgataatgataatg         | 4192 |
| Db | 4000 | AAAGTCAAGGAGTAAGTACTACCGAGATTAATAATTCGTCTTACCGAGATGTATGTAG         | 4059 |
| Qy | 4193 | gaattcccatctccagcagatcaaatggccaatgcaatggaactcttlttccaagaatggc      | 4252 |
| Db | 4060 | GAGTTCCTCATCTCCAGCAGCTAATGATGCGCATCTAGTGTGACTCTTTTTCAGAAATGCG      | 4119 |
| Qy | 4253 | aaagtaaaagttlttcaaggaatacaagactcctcacacctgtgtgataactctctagac       | 4312 |
| Db | 4120 | AAAGTAAAGGTTTTTCAGGGAATTAACAGCTCTTCAACACTGTGTGTAACTCTCTGAC         | 4179 |

QY 4313 ccaccgtactgactcgtctactctcgaattccacccccagagttgggtgaccagatggc 4372  
|||||  
Db 4180 CCACCGTACTGACTCGCTACTCTCGAATTACCCCAAGATTGGGTGCACAGATTGCC 4239  
QY 4373 ctgagatgagagttctggtgctgcagagcacagaccttac 4414  
|||||  
Db 4240 CTGAGATGAGGTTCTGGGCTGCGAGGACACAGACCTCTAC 4281

RESULT 5  
LOCUS 108642 108642 4548 bp PAT 02-DEC-1994  
DEFINITION Sequence 2 from Patent WO 8800831.  
ACCESSION 108642  
VERSION 108642.1 GI:588648  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4548)  
AUTHORS Unclassified.  
TITLE Pasek,M.P.  
DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR  
VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE  
POLYPEPTIDES IN HIGH YIELDS  
Patent: WO 8800831-A 2 11-FEB-1988:  
JOURNAL Location/Qualifiers  
FEATURES  
source 1..4548  
BASE COUNT 1344 a 1004 c 1002 g 1197 t 1 others  
ORIGIN

Query Match 92.5%; Score 4085; DB 10; Length 4548;  
Best Local Similarity 95.1%; Pred. No. 0;  
Matches 4325; Conservative 0; Mismatches 31; Indels 192; Gaps 3;

QY 59 gccaccagaatactactcgtggtgcagtggaactgtcactgagactatgcaagtgat 118  
|||||  
Db 1 GCCACCGAAGATACTACTCTGGTGACGTGAGTGTCAITGGGCTATATGCAAGATGAT 60  
QY 119 ctggtgagctgctctgtgagacgaagattcctcctcctagagtgccaaaactctttcaltc 178  
|||||  
Db 61 CTCGGTAGCGCTGCTGTGAGACGAATTCCTCTAGAGTGCACAAATCTTTCCATTTC 120  
QY 179 aacccttcagtcgtgtacaaaagacctctgtttgtaaatccaaggttccacctttcaac 238  
|||||  
Db 121 AACACCTCAGTCGTGTACAAAAGACCTCTGTTGTAGAAATTCACGATCACCTTTTCAC 180  
QY 239 atcgtaagccaagggcaccctggtgatgtctgtctagctccaccatccaggtcaggtc 298  
|||||  
Db 181 ATGCGTAAAGCAAGGCCACCTGATGGGTCTGCTAGTCTTACCATTCAGAGCTGAGGTT 240  
QY 299 tatgatacagtgatcaltactaactaagaacatggctccaccatcctgtcagttctcatgct 358  
|||||  
Db 241 TATGATACAGTGTGCTTATACCTTAAGACATGCTCCATCTGTCAGTCTTCATGCT 300  
QY 359 gttgggtatccctactcgtgaaagctctcgtgaaggagctgaataatgatagtatcagaccagtcaa 418  
|||||  
Db 301 GTTGGTATATCTTACTGGAAGCTTCTGAGGGAGCTGAATATGATGATCAGACACAGTCAA 360  
QY 419 agggagaagaagaatgataagcttcctcgtgtggaagccaataatgactgtcaggtc 478  
|||||  
Db 361 AGGGAGAAGAAGATGATTAAGTCTTCCTGTGTGGAAGCCATACATATGTCTGGCAGGTC 420  
QY 479 ctgaaaagaagaatggttccaatggcctcgtaccacactgtgcttactactacatactctct 538  
|||||  
Db 421 CTGAAAGAAGATGCTCAATGGCTCTGACCCACTGTGCTTACTACTCATATCTTTCT 480  
QY 539 catgtgagactcgtgtaaaagactgtaattcagggctcatttggagccctactagtatgtaga 598  
|||||  
Db 481 CATGTGAGCTGTTGTAAGAAGACTTGAATTGAGGGCTCATTTGAGAGCCCTACTAGATATGTAGA 540

QY 599 gaaggagtcgtgcgaaggaaagacacagaccttgcacaaatttactactcttggct 658  
|||||  
Db 541 GAAGGAGTCTTGGCCAAAGAAAACACACACCTTGGCAAAATTTATATCTTTTGGCT 600  
QY 659 gtaattgataaggaaagatttggcactcagaacaagaacctccttgaatgacaagataag 718  
|||||  
Db 601 GTATTTGATGAAGGAAAGTTGGCAGCTCAGAAACAAAGAACCTCTGATGACAGATAGG 660  
QY 719 gatgtgcattcgtctgggctcgtgctcaaaatgcacacagatcgaatgtaatgtataacagg 778  
|||||  
Db 661 GATGCTGATCTTGGCTGGGCTTGAATATGCAACAGTCAATGGTTATGTAAACAGG 720  
QY 779 tctctg---ccaagtcgattggaatgcccagagaaatcaagctatggaatgattgga 835  
|||||  
Db 721 TCTCTGCTACCAAGGTGATTTGATGATGCCACAGGAATACGTCTTTGGCATGTATTTGGA 780  
QY 836 atgggcaaccactccttgaagtgactcaataltcctcgaagggtcacaacttctgtgag 895  
|||||  
Db 781 ATGGCACACTCTCTGAAGTGCATCAATATTCCTCGAAGGTCAACATTTCTGTGAGG 840  
QY 896 aaccatgcgaagcgctccttggaaatctgcgaatcaatctcctactcgtccaacactc 955  
|||||  
Db 841 AACCATGCGCAGGGGCTCTTGAATATCGCCAAATTAATCTTACTGCTCAAAACATC 900  
QY 956 ttgattgaccttggacagattctactctgtttgtcatatctctccaccacaatgagtc 1015  
|||||  
Db 901 TTGATGACCTTGGACAGTCTTCTACTGTTTGTATATCTTCTCCACACATGATGGC 960  
QY 1016 atggaagcttatgtcacaagtgacagctgttccagaaggaaacccaactacgaatgaaat 1075  
|||||  
Db 961 ATGGAAGCTTATGTCAAGTATGACAGCTGTCACAGGACCCCAATACGAATGAAAAAT 1020  
QY 1076 aatgaagaagcggaaagactatgatatgactctactcattctgaatgatatgtgtcaag 1135  
|||||  
Db 1021 AATGAAGAAGCGGAAGACTATGATGATGATCTTCTGTAATTCGAATGAGATGCTGAGG 1080  
QY 1136 ttgatgatgacaactcctcctcctcttaccacaatctcgtcagttgcgaagaagactcct 1195  
|||||  
Db 1081 TTTGATGATGACAACTCTCTCTCTTATTCAAATTTGCTGCTAGTTGCCAAGAACATTCCT 1140  
QY 1196 aaaaacttgggtacatactaatgtcgtcgaagaaggagactggtgactatgctcccttagtc 1255  
|||||  
Db 1141 AAACTTGGGATTAATTCATTTGCTGTGAAGAGAGAGACTGGACTATGCTCCTTAGTC 1200  
QY 1256 ctggcccccagatgacagaagttataaagttcaatttgaacaatggccctcagcgatt 1315  
|||||  
Db 1201 CTCGCCCCCGATGACAGAAGTTATTAAGTCAATATTTGAACATYNNCCCTCACGGGATTT 1260  
QY 1316 gtaggaagtacaaaagaagtcgatttatagtacatacagaatgaaccttaagaactcgt 1375  
|||||  
Db 1261 GGTAGGAAGTACAAAAGTCCGATTTATGCTATACACAGATGAAACCTTTAAGACTTCGT 1320  
QY 1376 gaagctatccaagcaatgaatgaagaacttggaccttacttatttgggaagtgtgagac 1435  
|||||  
Db 1321 GAACCTATTCAGCATGATGATGAGAACTTGGGACTTTACTTTATGGGGAAGTTGGAGAC 1380  
QY 1436 acactgttgaattatatttaagaatccaagcagaagccatataacatctacccccaagga 1495  
|||||  
Db 1381 ACACCTGTTGATTAATTAATTAAGAAACAAGCAAGACGACCATATTAACATCTACCCACAGGA 1440  
QY 1496 atcacatgctcgtcctcttgaattcaagagagattacaaaaggtgtlaaacaatttgaag 1555  
|||||  
Db 1441 ATCACTATATGTCCTCTTTGTATTCAGAGAGATTTACCAAAAGGTGTAACAATTTGAAG 1500  
QY 1556 gatttccaattctgcagaggaataatccaataataatgagacgtgactgtagaagat 1615  
|||||  
Db 1501 GATTTTCCAATTTGCCAGAGGAATAATATTAATTAATGAAGACAGAGCTGTGAGAGAT 1560  
QY 1616 gggcacaactaaatcaagctcctcgtgtccgtgacccgctataactcttaattcgttaaatg 1675  
|||||  
Db 1561 GGGCAACTAATATCAATGATTCCTCGGTGCTGACCCGCTATTAATCTGTTAAATATG 1620  
QY 1676 gagagagatcagcttccagagctcattggtccctcctcctcactcgtcctaagaatctgta 1735

```
|||||
Db 1621 GAGGAGATCTAGCTTCAGGACTCATTTGGCCCTCTCTCATCTGCTACAAAGAACTCTGTA 1680
|||
Qy 1736 gatacaagaagaacccaataatgtcagacaagaagaatgtcatctgtttcttatt 1795
|||||
Db 1681 GATCAAAAGAGAAACCATATATGTGCACAAAGAGAAATGTATCTCTGTTTCTGTATT 1740
|||
Qy 1796 gatagaaccgaagctgtgtacctcacagaataatacaagcttctccccaatcagct 1855
|||||
Db 1741 GATGAGAACCAGAGCTGTACTCTACAGAGAAATATACAACGCTTCTCCCAATCCAGCT 1800
|||
Qy 1856 ggaatgcagcttgagatccagagttccaagcttccaacataatgacaagaatcaatgac 1915
|||||
Db 1801 GGAATGACGCTTGAGGATCCAGAGTTCCAAAGCCTCAACATCATCCACAGCATCAATGAC 1860
|||
Qy 1916 tatgttttgaatgttcagttgtcagttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1975
|||||
Db 1861 TATGTTTTGTAGTTTGCTACAGTTGTGATTTGTTGATGAGGTGGCATACTGGTACATT 1920
|||
Qy 1976 ctgaagcatgagacacagactgacttctctgtctctctctctctctctctctctctctct 2035
|||||
Db 1921 CTAGCATTTGGAGCAGACAGTGTACTTCTTCTGCTTCTTCTGATATACCTTCAAA 1980
|||
Qy 2036 cacaataatgtctatgaagaacactaaccttaccatcttcaaggagaatctgtcttc 2095
|||||
Db 1981 CACAAATGGCTATGAGAGACACACTCACTCCATTCCTCATCTCCAGAGAAACCTGCTTTC 2040
|||
Qy 2096 atgtcgaatggaacccccagctctatggaatctctgaggtgcccacaactcagacttgcgaac 2155
|||||
Db 2041 ATGTGATGAGAAACCCAGCTCTATGATTTCTGGGGTGCACACACTTCAGACTTTTGGAAAC 2100
|||
Qy 2156 agaggaatgacgcctctcctcgaaggttctcagttgtgacaagaacactggtgtattatc 2215
|||
Db 2101 AGAGAGCATGACCCGCTTCTGAAAGTTTCTAGTTGTGACAAAGAACACTGCTGATTATTAT 2160
|||
Qy 2216 gaggaacgttatgaagaatatttccagctacttgcgtgagtaaaaaaagttccattgaacca 2275
|||||
Db 2161 GAGGACAGTTATGAGAGATATTTTCACATATCTTGCTGAGTAAAAAACAATGCCATTGAACCA 2220
|||
Qy 2276 agaagcttctcccaagat- 2293
|||||
Db 2221 AGAAGCTTCTCCAGGATCTCTTGTGCGATTAACCACTATGTTACATACCAAAA 2280
|||
Qy 2294 - 2293
|||
Db 2281 GAAGATGGAATATCCAGAGAAAGTCAACAGAAAAACAGTTTAAAGAAAAAGATACC 2340
|||
Qy 2294 - 2293
|||
Db 2341 ATTTTGTCCCTGAACGCTTGTGAAGCAATCATGCAATAGCAGCAATTAATGAGGAGCAA 2400
|||
Qy 2294 - 2329
|||
Db 2401 AATTAAGCCCGAATAGAAAGTCACTGCGGCAAAAGCAAGTAGACATCGAAGAGCTGTGCTCT 2460
|||
Qy 2330 gccaccccaacagctctgnaaagcccaatcaacgaggaataactcgttactacttccagtca 2389
|||
Db 2461 CAAAAACCCACAGCTCTGAAAGCCCATCAACGGGCAAAATACCTGATCACTTCTAGTCA 2520
|||
Qy 2390 gatcaagagaagaattgactatgatatcaatcaatcagttgaaatgaagaagaagattt 2449
|||
Db 2521 GATTAAGAGAAATTTGACTATGATGATACCATATCATGTTGAATATAGAGAGAAATTTT 2580
|||
Qy 2450 gacatttatgataagatgaataatcaagaccccgcaacttccaagaagaacacagacac 2509
|||
Db 2581 GACTTTATATATGAGATGAGAAATCAGAGCCCGCAGCTTTCAAAAGAAAAACAGACAC 2640
|||
Qy 2510 tatttataatgctgagtgaggaaggtcttggaatataaggatgagtagtctccacatgt 2569
|||
Db 2641 TATTTATGTGCTGAGTGAGAGGCTTGAGATTTATGGATGAGTAGCTCCCCACATGTT 2700
|||
Qy 2570 ctgaagaacagaggtcagagtgagtgagtgctccctcagttccaagaagttgttttccagaa 2629
|||||
```

---

```
Db 2701 CTAAAGAACAGGGCTCAGAGTGGCAGTGTCCCTAGTTTCAAGAAAGTTGTTTCCAGGAA 2760
|||
Qy 2630 ttactatgagctcccttactcagccctataccgttgagagaactaaatgaactattgga 2689
|||||
Db 2761 TTTACTGATGAGTCTCTTACTAGCCCTTATACCGTGGAGAACTAATGAACATTTGGGA 2820
|||
Qy 2690 ctcctggggccatatataagagcagaagttgaagataatcatatgttaacttcaagaat 2749
|||
Db 2821 CTCCTGGGGCCATATATTAAGAGCAGAAAGTGAAGATATATCATATGTAATCTTCAAAAT 2880
|||
Qy 2750 cagagctctgcgcccatttctctatctatctagcccttatttcttcttgaggagaatcagaag 2809
|||||
Db 2881 CAGGCCCTCGCTCTATTCCTTATTCCTTATTCCTTATTCCTTATTCCTTATTCCTTATTCCT 2940
|||
Qy 2810 caaggaagagaaccagaaaaaacttgcagagccctaaatgaacccaacttacttctg 2869
|||||
Db 2941 CAAGAGAGCAAGAACCTAGAAAAAATCTTGTCAAGCCTATGAAACCAAACTTACTTTTGG 3000
|||
Qy 2870 aaagtgcacaatcatalatgagcaccccaataagaatgaagttgacgtgcaagccctgggttat 2929
|||||
Db 3001 AAAGTCAACATCATATGACACCCACTAAAGATGAGTTGACTCCAAAGCCTGGGCTTAT 3060
|||
Qy 2930 ttctctgattgtgaccttgaaaaaagtgtgcaactcagggccgatttgaaacctctgtgtc 2989
|||||
Db 3061 TTCTCTGATGTTGACCTGGAAAAAGATGTGACTGAGGCTGATTTGGACCTTCTGTGTC 3120
|||
Qy 2990 tgcacacttaacacactgaaacctgtctcatatgaggaagacagtgaaatgacagaaattgtct 3049
|||||
Db 3121 TGCCACACTAATACACACTGAAACCTGCTCATGSGAGACAAATGACAGTACAGAAATTTGCT 3180
|||
Qy 3050 ctgtttt---tcaacatcttgaatgaagccaaagctgtgtacttcaactgaataatgaa 3106
|||||
Db 3181 CTTGTTTTCCTCACCATCTTTGATGAGACCAAAAGCTGTCTTCTACAGAAATATGAA 3240
|||
Qy 3107 agaaactcagaggtccctctgcaatatccagaatgaagaatcccaattttaaagaagaattat 3166
|||
Db 3241 AAAAACTGAGGGCTCCCTGCAATATCCAGATGAGAGATCCCACTTTTAAAGAAATATAT 3300
|||
Qy 3167 cgcctccatgcatcaatgaatgactacataatgaatacaactcgtgttagtaatgagccag 3226
|||||
Db 3301 CGCTTCATGCAATCAATGAGGTACATATATGATGATACACTGCTGATTAATGAGTGCAG 3360
|||
Qy 3227 gatcaagaatctgaatgtatctgtcgaagcatgggcaagcaatgaataatcattatct 3286
|||||
Db 3361 GATCAAAAGATTCGATGATCTGCTAGCATGAGGACCAATGAATAATCATTCATTATT 3420
|||
Qy 3287 catctcagtgagaatgttcaatgttaacgttaaaaaaagaaggaggtataaattgagcattac 3346
|||||
Db 3421 CATTTCACTGACATGTGTCTCTCTGTACGAAAAAAGAGAGTATTAATATGGCACTGTAC 3480
|||
Qy 3347 aatcctcactcaggtgttttctgagacagtggaatgttaccatccaaagctggaatttg 3406
|||||
Db 3481 AATCTCTATCCAGGTGTTTGTGAGACAGTGAATAATGTATCATTCCAAAGCTGGAATTTGG 3540
|||
Qy 3407 cgggttggaatgcttataatgtggagagcatctacatgtctggatgagacaacttttctgtgt 3466
|||||
Db 3541 CGGGTGAATGCTTATTTGGAGAGCATCTACATCTCTGGGATGAGACACACTTTTCTGTGTG 3600
|||
Qy 3467 taacgcaataagtgctcagactccctctggaaatgtctcttgacacatlaagatcttcag 3526
|||||
Db 3601 TACAGCATTAAGTGTCAACTCTCCCTGGGAATGGCTTGTGACACATTAAGATTTTCAG 3660
|||
Qy 3527 attacagttcagaacaaatagacagtgggccccaagcttgccagactcattatctc 3586
|||||
Db 3661 ATTACAGCTTTCAGCAATATGAGACAGTGGGCCCAAAAGCTGGCAGACTTCATTATTTC 3720
|||
Qy 3587 ggaatcaatcaatgtcctgagacccaagagccctttcttctgtatccaaggtgagatctgtgt 3646
|||||
Db 3721 GGATCAATCAATGCTCTGAGACACCAAGAGGCCCTTTCTTGTGATCAAGGTGATCTGTG 3780
|||
Qy 3647 gcaacaatgattatctcagagcatlaaagcccaagggtgcccgttaagaagtttccaagctc 3706
|||||
Db 3781 GCACCAATGATTATTCAGGCAATCAAGCAACCAAGGAGTCCCTCAGAAAGTCTCCAGGCTC 3840
|||||
```

3707 tacatctcagttatcatcatatgatagtcttgatggaaagatggcagactatcga 3766  
|||||  
3841 TACATCTCTCAGTTATCATCATATATAGTCTTATATGGAAAGATGGCAGACTTATCGA 3900  
3767 ggaattccactggaaacttaattgctctctcttgcaatgltgatactcaggatataaa 3826  
|||||  
3901 GGAATTTCCACTGGAACTTAAATGCTCTCTTGGCAATGGATTCATCTGGGATFAAA 3960  
3827 cacaatttttaacctccaataatattgtccgataacatccglttgacccaactatcat 3886  
|||||  
3961 CACAAATTTTAACTCCCAATTTATGCTGATACATCCGTTTGCCCAACCACTTAT 4020  
3887 agaatcgcaagcaactcttcgatgagttgatggctgtaattaaatagttgagaatg 3946  
|||||  
4021 AGCATTTGCCAGCACTCTTCCGATGGAATTGATGGCTGTGATTTAAATAGTTGAGCATG 4080  
3947 ccattgggaatggagaaataagcaataatcagatgcacagataactgcttcatcttcc 4006  
|||||  
4081 CCATTTGGGAATGGAGATTAAGCAATATCAGATGCACAGATTACTGCTTACTCTT 4140  
4007 accaatattgtcgcacctggtctccttcaaaagctcgaactcaactccaaggagagat 4066  
|||||  
4141 ACCAAATATGTTTGCACCTGGTCTCTTCAAAAGCTGCACTTCACTCAAGGGAGAGT 4200  
4067 aatggcctggagagactcagtggaataatccaagaagatggtcgaagtggactccagaag 4126  
|||||  
4201 AATTCCTGGAGACCTCGGGAATTAATCAAAAGAGTGGCTGCAAGTGGACTTCCAGAG 4260  
4127 acaatgaataatcacaaggaatgaactactcagaaggaaataactctgcttaccagatgat 4186  
|||||  
4261 ACATGAAATATCAAGSAGTAATCACTACAGSAGTAATAATCTGCTTACCAGATGTAT 4320  
4187 gtaagaagttcctcatctccagcagatcaagaatgagcaatgaggaactccttcttcaag 4246  
|||||  
4321 GTGAGAGATTTCTCATCTCAGACAGTCAAGATGGCCATCAGTGAAGTCTCTTTTTCAG 4380  
4247 aatggcgaatgaaggtttctcagggaaatcaagaactccttcaactctggtggaaactct 4306  
|||||  
4381 AATGGCAAAATTAAGGTTTTCAGAGGAATCAAGACTCTTCACTGAGTGGTAAGTCT 4440  
4307 ctgaaccacacgttactgactcgtctactcgaatcaccaccagatltggtgacacag 4366  
|||||  
4441 CTGAGCCACCGTTACTGACTCGCTACTCTGCAATTCACCCCGAGATTGGTGGACACAG 4500  
4367 attgacctgagatgagatgtctggctgcgagggcacagaactctac 4414  
|||||  
4501 ATTGCCCTGAGATGAGGTTCTGGCTCGAGGACACAGGACCTCTAC 4548  
Db

RESULT 6  
LOCUS 108641 4551 bp PAT 02-DEC-1994  
DEFINITION Sequence 1 from Patent WO 8800831.  
ACCESSION 108641  
VERSION 108641.1 GI:588647  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 4551)  
AUTHORS Pasek,M.P.  
TITLE DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR  
POLYPEPTIDES IN HIGH YIELDS  
PATENT: WO 8800831-A 1 11-FEB-1988;  
JOURNAL location/Qualifiers  
FEATURES  
1..4551  
source /organism="unknown"  
BASE COUNT 1345 a 1003 c 1004 g 1198 t 1 others  
ORIGIN

Query Match 92.5%; Score 4085; DB 10; Length 4551;  
Best Local Similarity 95.1%; Pred. No. 0;  
Matches 4325; Conservative 0; Mismatches 31; Indels 192; Gaps 3;  
QY 59 gccacagaagataactactccttggtgacgttgaaactgtcatgtggaactatgtcaaatgtgat 118  
|||||  
Db 4 GCCACAGAGATATCACTGCTGGTGACAGTGAAGTGCATGAGGAGCTATATGCAAAAGAT 63  
119 ctgggtgagctgctgtggaacgaagattcctcctcgtagatgccaataatcttccattc 178  
|||||  
Db 64 CTGGGTGAGCTGCTGTGGAGCGCAAGATTTCCTCTGAGTGCACAAATCTTTTTCATTC 123  
QY 179 aacacccagtcgtgtgtaacaaagactctgtgtgaaattcaagttacagttcacttcaac 238  
|||||  
Db 124 AACACCTCAGTCTGTGTAACAAAGACTCTGTTTGTAGAATTACGGATACACCTTTTAAAC 183  
QY 239 atcgtaagccaagggccacacttgatgtgtcgtctgaagcttaccatccacagcggcgaagt 298  
|||||  
Db 184 ATGCTTAAGCCAAAGCCACCTGGATGGTCTGCTAGTCTTACATCCAGGCTGAGGTT 243  
QY 299 tatgatacagtggtcaatlaactaagaacatggcttccactcgtcagltcattcgtc 358  
|||||  
Db 244 TATGATACAGTGGTCATTACCTTAAGACATGGCTTCCATCTCTGACGTTCATGCT 303  
QY 359 gtgggtgatactacttgaaaagctctggaaggagctgaataatgtaagtaagcaagca 418  
|||||  
Db 304 GTTGTGATATCTTACTGTGAAAGCTTGTGAGGAGCTGAATATGATGATCAAGACCACTAA 363  
QY 419 agggagaagaagaatgaataagttcccttggttgaaagccaataatgtctggaagtc 478  
|||||  
Db 364 AGGAGAAAGATGATGAATGAAGTCTCCCTGTTGGAAGCCATACATATGCTGGCAAGTGC 423  
QY 479 ctgaagaagaatggtccaatggtccttgaccactggtccttaccatcatatcttctc 538  
|||||  
Db 424 CTGAAGAGATGATCCAAAGCTGCTGACCCACTGTGCTTACTATATCTTCT 483  
QY 539 catgtgagactggtgaagaagcttgaaattcagagccttcttgagccctcactcagatgtaga 598  
|||||  
Db 484 CATGTGACCTGCTGAAGACCTTGAATTCAGGCTCTATTGAGGCTTACTACTATGTAGA 543  
QY 599 gaaggagcttggtgcaagaagaagacacagactctgacaaattataactacttctgtc 658  
|||||  
Db 544 GAAGGAGTGTGGCCCAAGAAAGACACACACTTGCACAATTTATATCTTTTGGCT 603  
QY 659 gttattgataagggaaagtgttgcaactcagaagaacaaagactccttgatgagagataag 718  
|||||  
Db 604 GTATTGTATGAAGGAAAGTGTGGCACTTCAGAAACAAAGAACTCTTGATGACAGATAGG 663  
QY 719 gatgctgcatctgcttggtggtcgtgacctaaatgacaagtcgaatggttataagacagg 778  
|||||  
Db 664 GATGCTCATCTGCTCGGCTGAGCTTAAATGCAACACAGTCAATGGTTATGTTAAACAGG 723  
QY 779 tctctg---ccaggtctgattgtagtgcacaggaatcaatctatltgacatgtgatgga 835  
|||||  
Db 724 TCTCTGTACAGAGTGTGATGTGATGTCACAGGAATCAATGTATGATGATGGA 783  
QY 836 atgggacacactccttgatgacatcaatattcctcgaaggtcaacacattctgtgagg 895  
|||||  
Db 784 ATGGGCACTCCTCTGAAGTGCACCTCAATATCTCTCAAGGTCAACACATTTCTTGAGG 843  
QY 896 aacacatcgcaagcgtctcgtggaatcgcacataactcttccattactgttcaaaactc 955  
|||||  
Db 844 AACCATGCGCAGGCGTCTTGGAAATTCGCCAATATCTTCTTACTGCTCAAAACACTC 903  
QY 956 ttgatgaccttgagacgttcttactgttctgtcatalactcttccaccacacatgtatgac 1015  
|||||  
Db 904 TTGATGACCTTGAGACGTTTACTGTTTGTGATATCTCTTCCCAACCAATGATGAGC 963  
QY 1016 atggaagcttatgtcaaatgataagacgtgtccagaggaacccaactagaatgaataat 1075  
|||||  
Db 964 ATGGAAGCTTATGTCAAAAGTGAACAGCTGTCCAGAGGAACCCCAACATGCAAAAT 1023  
QY 1076 aatgaagaagcgggaagactatgtatgatcttactcattcgaatgagtggtgacagg 1135

|    |      |  |      |
|----|------|--|------|
| Db | 1024 | AAATGAAGAACCGGAGACATGATGATGATCTTACTGATTTCTGAAATGGATGTGTCCAGG       | 1083 |
| Qy | 1136 | tttggatgatgacaactctccctcccttataccaatcgcctcagttcgtccaaagaagatcct    | 1195 |
| Db | 1084 | TTTGGATGATGACAACTCTCCTCTCTTATTCCAAATGTGCTCAGTTGCCAAGACATCTCT       | 1143 |
| Qy | 1196 | aaaacttgggtacattacattcgtctgctgaaaggaagagcctggtgactatgctcccttagtc   | 1255 |
| Db | 1144 | AAAACTTGGGTGACTATTCATTTGCTGCTGAAGAGGAGAGCTGGAGCTATAGTCCCTTACTGC    | 1203 |
| Qy | 1256 | ctgcgccccgcatgacagaagttataaaagtcatttttgaaaccaatggcccccaagsgaatt    | 1315 |
| Db | 1204 | CTCGCCCCGATGACAAAGTTATTAAGTCAATATTTTGAACAAATGGCCCTCAGGGGATTT       | 1263 |
| Qy | 1316 | ggtgaaggaagacaaaagaatcccgatcttatacgcatcacacagatgaagaacttaagaactcgt | 1375 |
| Db | 1264 | GGTAGGAAGACACAAAAAGTCCGATTTATGCGATACACAGATGGAACCTTTTAAGACTCGT      | 1323 |
| Qy | 1376 | gaagctattcaagcatgataccaagaatcttggagcccttacttatactggggaagtctggagac  | 1435 |
| Db | 1324 | GAACTATTTCAGCATGATCAATCAGGAATCTTGAGACTTTACTTTATTTAGGGAAGTTGGAGAC   | 1383 |
| Qy | 1436 | acacgtttgtattatatttaagaatccaagcaagcagacacatatatacatctacccctcaga    | 1495 |
| Db | 1384 | ACACTGTGATTTATTTAAGAAATCAAGCAAGACAGACCATATTAACATCTACCCCTCAGCA      | 1443 |
| Qy | 1496 | atcacggaatgcgcctcttctgattccaaggaatctaccaagaagttgaacaacttgaag       | 1555 |
| Db | 1444 | ATCACATGATGTCCGCTCTTTGTATTCAAGGAGATTTACCAAAAAGTGTAACATTTGAAAG      | 1503 |
| Qy | 1556 | gatttcccaattctctgcacaggaagaatactcaaatataaattgagacagtcgttagaagat    | 1615 |
| Db | 1504 | GATTTTCCAAATCTGCGCAGGAGAAATATTCAAATATATAATGAGACAGTACGTAGTAAGAT     | 1563 |
| Qy | 1616 | ggggccaactcaatcagaatcctctgggtgccttgaccccgctatactctagtttcgttaaatg   | 1675 |
| Db | 1564 | GGGCCAACTAAATCAGATCTCGGTGCTGACCCGCTATTAATCTGTAGTTTGTTAATATG        | 1623 |
| Qy | 1676 | gaagaagatctagcttcaaggaatcatctggccctctccatctctgatacaagaatctgta      | 1735 |
| Db | 1624 | GAGAGAGATTTAGCTTTAGAGATCATTTGGCCCTCTCTCATCTGCTACAAAGAAATCTGTA      | 1683 |
| Qy | 1736 | gatcaagaagaaaccaagataatgttcagacaagaagaaatgctacatcctgttttctgtatt    | 1795 |
| Db | 1684 | GATCAAGAAGAGAAACCGATATATGTCAGACAAAGAGAAATGTCATCTGTTTCTGTATTT       | 1743 |
| Qy | 1796 | gatgagaaccgaagctggtatccttcaacagagaatatacaagcttcttcccaatccagct      | 1855 |
| Db | 1744 | GAGGAAGAACCGAAGCTGTGTACTCACAGAGAAATATACAACTTTTCTCCCAATCCACACT      | 1803 |
| Qy | 1856 | ggagtgcaagtttgaatgtccagagttccaagcctccaacatcatgacaagcatcaatggc      | 1915 |
| Db | 1804 | GGAGTCAAGCTTGAAGATCCAGAGTTCCAGAGCTTCACACATCATGCAACATCANTATGCG      | 1863 |
| Qy | 1916 | tatgttttttgaatgttgcagttgttcagttgttgcattgagatggtggcactactgtgtacatt  | 1975 |
| Db | 1864 | TATGTTTTTGAATGTTTGGAGTTGTGAGTTTGTTTGCAATGAGTGGCATTAATCTGTAAT       | 1923 |
| Qy | 1976 | ctaagcatctggagacaacagctgacttccctctgtctctctctctgtatataccctcaaa      | 2035 |
| Db | 1924 | CTTAAGCATTTGGACACACAGACTGACTTCTTCTTCTCTCTCTCTGATATACCTTCAAA        | 1983 |
| Qy | 2036 | caaaaaatgtctcatgaagaacacactaacctcatctccatcttcaagagaacgtctctc       | 2095 |
| Db | 1984 | CACAAAATGTCTATATGAAGACACACTCACTCACTATTTCCATTTCTCAGGAGAAATGTCTTTC   | 2043 |
| Qy | 2096 | atgtcgaatgaaaaaccaggatctiatgtatcttgggtgtccaaactcagacttccggagac     | 2155 |
| Db | 2044 | ATGTCTCATGAAAAACCCAGGCTTATGCAATTTCTGGGGTGCACAACTCAGACTTTTCGGAAC    | 2103 |
| Qy | 2156 | agagggcatgacgcgcttactcgaaggttcttctagttgttgacaagaacactggtgatattac   | 2215 |
| Db | 2104 | AGAGCATGACCGCCTTACTGAGGTTTCTAGTTGTGACAAAGAACACTGGATTTATTAC         | 2163 |
| Qy | 2216 | gaggaacattatgaaagatalttcaagcatactgtctgagtaaaaaacatlyccatttgaacca   | 2275 |
| Db | 2164 | GAGGACAGCTTATGAAGATATTTTACACATACATTGCTGACTTAAACAAATGCCATTGAACCA    | 2223 |
| Qy | 2276 | agaagcttctcccaaat-----   | 2293 |
| Db | 2224 | AGAACTTCTCCAGAGATCTCCTTTCCTTGGATTAACACACTATGTTACTCAGATACCAAA       | 2283 |
| Qy | 2294 | -----  | 2293 |
| Db | 2284 | GAAAGTGGAAATCCCAAGAGAAAGTCACACAAAAAACAGCTTTTAAGAAAAAGGATACC        | 2343 |
| Qy | 2294 | -----  | 2293 |
| Db | 2344 | ATTTTGTCCCTGAACGTTGTGAAGCAATCATGCAATAGCAGCAATTAATGAGGACAA          | 2403 |
| Qy | 2294 | -----  | 2329 |
| Db | 2404 | AATAAGCCCGAAATGAAAGTCACTGGGCAAGGCAAGGTAGACAGTGAAGGCTGTGCTCT        | 2463 |
| Qy | 2330 | gcccccacacagctcttgaagaagccataaagggaataaactctgtactactcttcagtca      | 2389 |
| Db | 2464 | CAAAACCCACAGCTTTGAAGCGCATPAAGGGGAATTAACCTCGTACTACTCTTCAGTCA        | 2533 |
| Qy | 2390 | galcaagaaggaattgactatgatataccataltaacttgaattgaagaagaagaatttt       | 2449 |
| Db | 2524 | GATCAAGAGGAATGACTATGATGATATACATATATCAGTTGAATGAATGAACAGGAATTTT      | 2583 |
| Qy | 2450 | gacatttatgtagagatgaataatcagaagcccccgacgttctcaaaagaanaacagacac      | 2509 |
| Db | 2584 | GACATTTATGATGATGAGAGAAATCAGAGCCCGCACCTTTCAAAAAGAAACAGACAC          | 2643 |
| Qy | 2510 | tatttatttgcgcagtggaagagctcgggattatggatggatagctccccaatglt           | 2569 |
| Db | 2644 | TATTTTATTTGCTGAGTGAAGAGGCTCTGGGATTTATGGAGTGAAGTCCCACTGTT           | 2703 |
| Qy | 2570 | ctaaagaacagggctcagaagtggcagtgctccctcagttccaagaagttgttttccagaa      | 2629 |
| Db | 2704 | CTAAGAAACAGGCTCAAGATGGCGAGTGTCCCTCAGTTCAAGAAAGTTGTTTCCAGSAA        | 2763 |
| Qy | 2630 | ttactgatgagctcttactcactcaagcccttataccgttggagaactaaatgaaatctggga    | 2689 |
| Db | 2764 | TTTACTGATGGCTCCTTACTACAGCCCTTATTAACCTGAGAACTTAATGAACATTTGGA        | 2823 |
| Qy | 2690 | ctccctggggccatatataagaagcagaagtgtgaataatatacatgataacttcaagaat      | 2749 |
| Db | 2824 | CTCCTGGGGCCATATATTAAGACCAAGATTTGAAGTATATATCATGTACTTTTCAAGAT        | 2883 |
| Qy | 2750 | caggcctctgctccctattcctctcatctagcctctattcttcttgaagaagatccaag        | 2809 |
| Db | 2884 | CAGGCTCTCGTCCCTATTCTCTTATCTTAGCTTATTTCTTATGAAGAGATCAAGAG           | 2943 |
| Qy | 2810 | caaggaagcagaacctagaaaaaacttgtccaagccctaaatgaacocaaacttactttgg      | 2869 |
| Db | 2944 | CAAGGAGCAGAACTTGAAGAAACTTTGTCAAGCCTATGAAACCAAACTTACTTTTGG          | 3003 |
| Qy | 2870 | aaagtgcacaatcatatgtgcacccactaaagaatgagattgtgactgcaagccttgggtctat   | 2929 |
| Db | 3004 | AAAGTGCAAATCATATGTGCAACCTTAAGATGATGTTTGACTGCTCAAAAGCTTGGGCTTAT     | 3063 |
| Qy | 2930 | ttctctgattgtgaccttgaaaaaagaatgtgacatcaagcctgcatgtgagacccctctgctc   | 2989 |
| Db | 3064 | TTCTTGATGTGACCTTGAAAAAGATGTGCACTCAGGCTTGATTTGGACCCCTTCTGTGTC       | 3123 |
| Qy | 2990 | tgccacacatacacactigaaccctgtctatgtggagacaaagtgaacagttaacagaaatttcgt | 3049 |
| Db | 3124 | TGCCACACTTAACACACTGAACCTGTCTATGTGGAGACAAAGTGAACAGATTTTGCT          | 3183 |
| Qy | 3050 | ctglttt---tcaaccatcttgatgagacaagaagctggttacttcaactgaaataatgaa      | 3106 |
| Db | 3184 | CTGTTTTCCTCACCATCTTTGATGAGACCAAAAGCTGTGCTTCTCAGTGAATAATGTGAA       | 3243 |



3107 agaaactgcagggtccctcgaataaccagatggaagatcccaattttaaagaagaattat 3166  
|||||  
3244 AGAAACTGCAGGGCTCCCTGCATATCCAGATGSAAGATCCACTTTTAAAGAAATTAAT 3303  
3167 cgttcacatgaatcaatggtctacataatgatatactactggtctgtagtaatggtccag 3226  
|||||  
3304 CGCTTCATGCAATCAATGGGTACATTAATGGATACACTACTGGCTTAGTAATGNTCAG 3363  
3227 gatcaaaagattcgatgtagtctcgtcgaagatgggcaaatgaaacatccattctattc 3286  
3364 GATCAAAAGGATTCGATGATCTCTCAGCATGGGCAATGAAACATCCATTCATCTATT 3423  
3287 cattcagtgagacatggttcacatgtaaaaaaagaagaagataaaatgggcacttac 3346  
3424 CATTTCAGTGACATGTTGTTTCTACTGTACGAAAAAAGAGATTAATAATGGACTGTAC 3483  
3347 aatctctatccagggtgtcttttgagacagtggaatgttacatccaaagcttgaaatttg 3406  
|||||  
3484 AATCTATATCCAGGTTGTTTGAAGACAGTGAATGTTACCATCCAAAGCTGGAATTTGG 3543  
3407 cgggtggaatgcttatttggcgaagatctacatgctggatgagacatttctgttg 3466  
3544 CGGGTGAATGCCCTTATTTGGCAGACATCTACATGCTGGATGACACACTTTTCTGCTG 3603  
3467 tacagcaataagtgtagactcccttggaatggtctcttgacacattagaagatttcag 3526  
|||||  
3604 TACAGCATATAGTGTGAGACTCCCTGGAAATGGCTTCTGACACATTAGGATTTTCAG 3663  
3527 attacagcttcaggaacaataatgagacagtgggcccaaaagcttgccagacttcatlcc 3586  
3664 ATTAACAGCTTCAGAGCAATATGACAGTGGGCCCAAAAGCTGGCCACACTTCATTATTC 3723  
3587 ggtatcaatcaatgctcggagagacaaagagaccccttcttgatcaaatggtggtatctg 3646  
3724 GGATCAATCATGCTGAGACACCAAGAGACCCCTTTCTTGATCAAGGTGATCTGTTG 3783  
3647 gcaaccaatgatttaacagcagcatcaaaagcgaagtgccgctcaagaagtcttccagctc 3706  
|||||  
3784 GCACCAATGATTTATTCACGGCATCAAGACCAAGGTCCTCCGTCAGAACTTCTCCAGCTTC 3843  
3707 tacactctcagttatcalcatgtagttagtcttgatgagaagaatggtcagactatcga 3766  
3844 TACATCTCTCAGTTTATCATCATGATAGTCTGATGGAAGAGTGGCAGACTTATTCGA 3903  
3767 ggaatctccactggaacttaatggtctcttgggcaatggtgattcaatctgggataaa 3826  
3904 GGAATTCACCTGGAACTTAATGCTTCTTTGGCAATGTGATCATCTGGCATTAATAA 3963  
3827 cacaatttttaacctccaattatgtcgaatacgcgttgcacccaactcatat 3886  
3964 CACAATTTTTTAACCTCCAAATTAATGCTCGATACATCCGTTTGCCCAACTCATTAAT 4023  
3887 agcatctcagcactcttcgcatggaatgtagtggctgtagtattaaatgltcgaagatg 3946  
4024 AGCATTCGACACTCTTCCATGAGAGTGTATGGGCTGTGATTTAATAATGTTGAGCATG 4083  
3947 ccatttgggaatggaagataagcaataatcagaatgacagattactgcttcaatccattc 4006  
4084 CCAATGGGAATGAGATTAAGCAATATCAGATGCACAGATTAATGCTTCATCTACTTTT 4143  
4007 accaatatgttgcacctggtctcccttcaaaagctgacttcaacctccaaggaagaagt 4066  
4144 ACCAATATGTTTGGCCACTGTGTCTCTTTCAAAAGCTGCACTTCACCTCCAAAGGAGAGT 4203  
4067 aatgctcgaagacctcagtgtaataatccaaagaatggtcgtcgaagtgcgttccagaag 4126  
4204 AATGCTCGAGACTCGATGGAATTAATCCAAAGAGTGGCTGCAAGTGGCACTTCCAGAG 4263  
4127 acaatggaatcacaaggaatgaactactcgaaggatlaaatctctgcttaccagatgat 4186  
4264 ACAATGAAGTCAAGAGATTAATCACTACAGGAGGATTAATAATCTGCTTACCAGCATGTAT 4323

4187 gtaaggaattccctcatctccagcagatgagatgcatcagtgactctcttcttcag 4246  
|||||  
4324 GTGAAGAGATTCTTCATCTCCAGCAGTCAGATGSCCATCATGATGACTCTTTTTCAG 4383  
4247 aatgcaaaatlaaagtttctcaaggaaatcaaatcactcttcaacctgtgtgaactct 4306  
4384 AATGCAAAAGTAAAGTTTTTCAGGAAATCAAGACTCTTTCACACCTGTGTGTAACCT 4443  
4307 ctgaaccacacgttactgactgctactccttcgaatttcaaccccaagattgggtgacag 4366  
4444 CTGACACCCACCGTTTACGACTGCTACTTTCGAATTCACCCCGAAGTTGGGTGCACAG 4503  
4367 atgcccctgaagatggaagtctggtcgtcgagagacagaacctctac 4414  
4504 ATTGCCCTGAGATGAGAGTTCTGGCTGCGAGGACACAGACCTCTAC 4551

RESULT 7  
AR034084 5035 bp DNA PAT 29-SEP-1999  
LOCUS Sequence 1 from patent US 5869292.  
DEFINITION  
AR034084  
ACCESSION  
AR034084.1 GI:5949689  
VERSION  
AR034084.1  
KEYWORDS  
Unknown.  
SOURCE  
Unknown.  
ORGANISM  
Unclassified.  
REFERENCE 1 (bases 1 to 5035)  
AUTHORS Voorberg,J.J.  
TITLE Hybrid proteins with modified activity  
JOURNAL Patent: US 5869292-A 1 09-FEB-1999;  
FEATURES  
source 1..5035  
BASE COUNT 1484 a 1127 c 1110 g 1314 t  
ORIGIN

Query Match 84.8%; Score 3743.8; DB 9; Length 5035;  
Best Local Similarity 87.5%; Pred. No. 0;  
Matches 4365; Conservative 0; Mismatches 52; Indels 570; Gaps 2;

41 catgcaaatagagctcccaactgctctctctgtgcttcttgatctcgtcttaagtc 60  
|||||  
34 CATGGAATAGAGCTCTCCACTGCTTCTTCTGTGCTTGTGGATTCCTTTAGTGC 93  
61 caccagaagatcactcctggtgagtgagacgtgcatggaactatgagatataatgcaagtatct 120  
94 CACCAGAAGATACCTGAGTGGTGCAGTGAATGTCAATGAGACTATATGCAAACTGATCT 153  
121 cgggtgagctgctgtgagcgaagatttccctcagaagtgcacaaatcttccattcaa 180  
154 CGGTGACCTGCTGTGAGACGCAAGATTTCTCTAGAGTGCAAAATCTTTTCATTCAA 213  
181 caactagctggtgacaanaaagactggtttagaattcaacgttcaaccttcaaat 240  
214 CACCTCAGTGTGTACAAAAAGACTGTGTGTGAATTCAGGATACACTTTTCAACAT 273  
241 cgttaagcacaagccacctggaatggtctgtagtgccttcaatccaaagctggaagtta 300  
274 CGCTTAAGCCAAAGGCCACCTGAGATGGGTGCTGATGCTCAATCCAGGCTGAGGTTTA 333  
301 tgatagagtggttaatttaagaacatggtcttcccatccctgtagtcttcttctgtc 360  
334 TGATACAGTGTCTATTAACCTTAAGAACATGGCTTCCATCTCTGCTGAGTGTG 393  
361 tggtagtctcactggaagctcttgaggagctggaatgatatgatatgaagccaatcaag 420  
394 TGGGTATCTCTACTCGGAAGCTTGTGAGGAGCTGGAATATGATGATACACCACTCAAG 453  
421 ggaagaaagatgataaagltctccctggtggaagccatacatatgctcgtgaagtcct 480  
454 GGAGAAAGAGATGATTAAGTCTTCCCTGGTGAAGCCATATATATGTGGAGGTCTT 513



481 gaaagaaatggtccaaatgacctgacccactgcttacctactcaatcattcttca 540  
|||||  
514 GAAAGAAATGTCCTCAATGGGCTGTGACCACTGTGCTTACTCTACTATCTTTCTCA 573  
541 tctgacctggttaaaagacttgaatcaggccctcatttggaccctactagtgtagaga 600  
|||||  
574 TGTGGACCTGTGTAAAGACTGGAATTCAGGGCTCATTTGGAGCCACTAGTGTAGAGA 633  
601 agggagcttgcccaaggaaaagacacagaccttgacaaaattactactcttctgctg 660  
|||||  
634 AGGAGTCTGGCCAGAGAAAGACACAGACCTTGCCACAAATTTATCTACTTTTGCTGT 693  
661 atttgatgaagggaanaatttggcactcagaacaaagaactccttgatgcaagatagga 720  
694 ATTGTGATGAAGGGAANAATTGGCACTCAGAAACAAAGAACTCTTGATGACAGATAGGA 753  
721 tctgcaatctgctcgagcctgacctaaatgcaacagtcacagtcaatggttatgtaacagtc 780  
|||||  
754 TGTGCAATCTGCTGGGGCTTGCCCTTAAATGCACAGTCAATGGTTATGTAAACAGTTC 813  
781 tctgcaagctctgatttggatgcccacaggaataatcagtcatacttggcaatgtaattggatgg 840  
|||||  
814 TCTGCCAGGTGTGATTTGATGCCCACAGAAATCACTTATTTGGCATGTGATTTGGATGGG 873  
841 caaccactctgaagtgcaactcaatacttcgaaagtgcaacacattcttgttgaaggaacca 900  
|||||  
874 CACCACCTCTGAAGTGCACTCAATATTCCTCGAAGTGCAACATTTCTGTGAGAGACA 933  
901 tctgcaagcgtctcttggaaatctcgcaataaacttctccttactgctcaaaccttctgat 960  
934 TCGGCAGGCGCTCTGGAAATCTGCCCAATTAATCTTCTACTGCTCAAAACCTCTTGAT 993  
961 ggaaccttggaaagtttctactgcttctgcatatctcttcccaacaaactgataagcatgga 1020  
994 GGACCTTGAGACAGTTTCTACTGTTTGTGATATCTCTTCCACCAACATGATGGCATGGA 1053  
1021 agcttactgcaaaagttagacaagctgtccagaaggaaccccaactactagaatgaanaataatga 1080  
1054 AGCTTATGTCAAAAGTAGACAGCTGTCCAGAGAAACCCCACTAGCAATGAANAATAATGA 1113  
1081 agaaagggaagaactactagatgatacttactgatactcgaatgagatggttgaatttga 1140  
1114 AGAAGCGGAACACTATGATGATCTTATGATCTTGAATGATGATGCTGAGCTTTGA 1173  
1141 tgaatgaacaactctccttcccttacttccaaatccgctcagtttgcagaagaagcatcccaaac 1200  
1174 TGAATGACAACCTCTCTCTTCTTATTCACAAATTCGCTCACTTGCCAAAGACATCTTAAAC 1233  
1201 tctggatacatatacttgcgtctgaaaggagagacttggactatgctcccttagtctcgcg 1260  
1234 TTGGGTACATTAACATGCTGCTGAAGAGAGAGACTGAGACTATGCTCCCTTATGTCCTGCG 1293  
1261 ccccgatgaagaagattataaagttcaatatttgaacaatggccctcagcgatggttag 1320  
1294 CCCCAGTAGACGAAGATTATAAAGTCAATATTTGAACAATGGCCCTCGAGGGATTGGTAG 1353  
1321 gaagtacaaaaaagctcgacttatactgcaatacagaatgaacacttaagaactcgatgaagc 1380  
1354 GAAGTACAAAAAAGCTCGATTATATGGCATACAGATGAACCTTTAAGACTCGTGAAGC 1413  
1381 tattcagcatgaatcagaatcttgggaacttacttatagggaagtttggagacacact 1440  
1414 TATTCACATGAATCAGGAATCTTGGGACCTTACTTTATGGGGAAGTTGGAGACACT 1473  
1441 gttgatatataatgaagaatcaagaacagacacataataacatctacccctcacygaataac 1500  
1474 GTTGATTATATTATTAAGATCAAGCAAGCAGACATATTAACATCTTCCCTCAGGAGATCAC 1533  
1501 tgaatctgctccttctgtatcaagaagattacaaaagggttaaaaacttgaagaattt 1560  
1534 TGAATGCTCGCTCTTTGTATTCAAGGATTTACCAAAAGGTGTAANAACATTTGAAGGATTT 1593

1561 tccaaatcttccaggagaanaatttcaaatataatlgacaagtgactgaagaatgggccc 1620  
|||||  
1594 TCCAAATTTGCCAGAGAAATATTCAAAATATTAATGACAGTACTAGAAAGATGGGCC 1653  
1621 aactaaatcagatctctgggtgctgtaaccgctattactctagtgttctgttaatatggagag 1680  
|||||  
1654 AACTAATACATGATCTCGGTGCTTACCCGCTATTTACTGTGTGTTCTGTAATATGGAGAG 1713  
1681 agatctagcttcaagactcaatctggccctctccctcactcgtcacaagaactcgttagatca 1740  
1714 AGATCTAGCTTCAAGACTCATCTGAGCCCTCTCTCTCATCTGCTACAAAGATCTGTAGATCA 1773  
1741 aaggagaaaccagataatgctcagaagaaggaatgtcattcctgttcttctgtatttga 1800  
1774 AAGAGAAACCCAGATATATGTAGACAAAGAGATGTCTCTGTTTCTGTATTTGATGA 1833  
1801 gaaccggaagctgtgtatccctcacaagaagaataataaagcttcttccccaatcgaagtcagtc 1860  
1834 GAACCGAAGCTGTGTACTGTACAGAGAAATATCAACGCTTTCTCCCAATTCACGCTGAGT 1893  
1861 gcaacttgaggaatccagaagtttccagctccacaatactgcaagacatcaatggctatgt 1920  
1894 GCAGCTTGAGATTCAGAGTTTCCAGCTCCAACTCATGTGACAGACATCATGTGCTATGT 1953  
1921 ttgtatagtttgcagttgtcagtttgttgcataaggtggcaactgtgtacatcttaag 1980  
1954 TTTGTATGTTTGGAGTTGTCTAGTTGTGATGATGAGTGAGTGCATCTTCAATCTTAAG 2013  
1981 catitgagcagaagacttcccttctgtcttctctctctgatatactctcaaacaa 2040  
2014 CATTTGAGACACAGACTGTACTCTTCTGTCTCTCTCTGATATATCTTCAACACAA 2073  
2041 aatgttctatgaagacacactcaccctatcccaatctcagaagaanaactgtctcactgtc 2100  
2074 AATGCTTATGATGAAGACACACTACCCCTATTCCTCATCTCAGAGAAATGCTTCACTGTC 2133  
2101 gatggaaaaaccaggtctatgattcttggggttgcacaactcgaacttctggaaacagag 2160  
2134 GATGAAAAACCCAGGTATGATGATCTGGGGTGCACAACTCAGACTTTCGGAACAGAGG 2193  
2161 catgacgcttactgaaaggttcttctagtgtgacaagaacacgtgtgtattctacgggga 2220  
2194 CATGACCGCTTACTGGAAGGTTTCTAATTGATTTCCAGAGGGGAGGAGCGACGACTA 2253  
2221 cagtatgaatattca-----gcaatcttctgagttaaacaatgctc 2265  
2254 TCTGCACTGTGAGATATTCAGTGAAGAGAGAGACTACATGACATCTGCGACACTCT 2313  
2266 catgaaaccaagaagctcttccagaatctcaagaacccctagcaactgagcagaagaactc 2325  
2314 GATTGAACCAAGAACCTTCTCCCAAAATTCAAAGACACCCCTAGCACTAGGCCAANAAGCAATT 2373  
2326 taatgcaac----- 2334  
2374 TATATGCCACCAATTTCCAGAAATGACATAGAGAAAGTCAACCTTGGTTTGCACACAG 2433  
2335 ----- 2334  
2434 AACACCTATGCTTAATAATACAAATGTCTGCTAGTATTTGATGCTCTTGCGACA 2493  
2494 GAGTCTACTCCACATGGGCTATCCCTTATCTGATCTCCMAAGCAAAATATGAGACTTT 2553  
2335 ----- 2334  
2534 TTCTGATGATCATCACTGTGAGCAATAGACAGTAATTAACAGCTGTCTGAAATGACACA 2613  
2335 ----- 2334  
2614 CTTTCAAGGCACAGCTTCATCAGAGTGGGACATGATTTACCCCTGATGACGACTTCCA 2673  
2335 ----- 2334

|    |      |  |       |
|----|------|--|-------|
| Dd | 2674 | ATTAAAGATTTAATGAGAACAACCTGGGGCAACACTGCAGATTCCTTGTGGATTAAACCATA     | 27333 |
| Oy | 2335 | -----  | 2334  |
| Dd | 2734 | TGGTAGTCAGATTAACCAAACGAAGTGGAATGCCAAGGAAGTCCACGAAAAACAGC           | 27333 |
| Oy | 2335 | -----  | 2334  |
| Dd | 2794 | TTTTAAAGAAAAGGATACATTTTGTCCCGTGAACGCTGTGAAGAACCATATCAATAGCATAGC    | 28533 |
| Oy | 2335 | -----  | 2334  |
| Dd | 2854 | AGCAATAAATGAGGAGCACAAATAAGCCCGAATAATGAAAGTCACCTGGGCAAGCAAGTAG      | 29133 |
| Oy | 2335 | -----ccacaagcttggaaagcga tcaagggaaataac<br>                        | 2370  |
| Dd | 2914 | GACGTGAAGGCGTGTGCTCTAAAAACCCACAGCTTGAAGACCCTATCAACGGGAATATAC       | 29733 |
| Oy | 2371 | tcgtaactcttcogtlcaga tcaagaggaatltgactlgtatlaccatcacgtlgtga        | 2430  |
| Dd | 2974 | TCGTACTACTCTTCAGTCAGATCATGAAGGAATAATTGACTATGATGTATACCATATCAATTGTA  | 30333 |
| Oy | 2431 | aabtaaagaaagaattttgcacatltaibaitagatlganaalcaagcccccaagt           | 2490  |
| Dd | 3034 | AATGAAGAGGAAGATTTTGCATTTATATATGATGAGTGAATAATCAGAGCCCCGAGGTT        | 30933 |
| Oy | 2491 | tcaaagaagaacacgaacactlltta ltblgtcbagtgtagagcctctygga tlatlygat    | 2550  |
| Dd | 3094 | TCAAAAGAAAACACGACACTATTTTATGTGCTGAGTGAGGAGAGGCTCTGGATTAATGAGAT     | 31533 |
| Oy | 2551 | gaftagctcccacacigtltctaagaaacagagctcagagtygcagtlyrccctcautlcaa     | 2610  |
| Dd | 3154 | GAGTAGCTCCCACATGTGTTCTAAGAAAACAGGGCTCAGATGGCAAGTGCCTCAGTTCCA       | 32133 |
| Oy | 2611 | gaagtggttttccaagaaatltaactblgtbcctcttactcagacccttatccgltaga        | 2670  |
| Dd | 3214 | GAAGTGTGTTTCCAGGAATTTACTATGATGGCTCTTACTCAGCCCTTATACCTGGAGGA        | 32733 |
| Oy | 2671 | actaaatgaacatltyggacctcclyggscatatalaagagcagaagittgaagataaat       | 27330 |
| Dd | 3274 | ACTAAATGAACATTTTGGAGACTCCTGGGCCATATATAAGCAGAAAGTTGAAGATAATAT       | 33333 |
| Oy | 2731 | cabgtacttctaagaatcagggccctctgtccatctccctctatctctatcttbaacctatitc   | 27930 |
| Dd | 3334 | CATGTACTCTTCAAGAATATAGGSCCTCTGGCTCCATTCTCTTATVCTTACCTTTATTTTC      | 33933 |
| Oy | 2791 | ltaiaggaagaatlcagagcaagagcagaacactgaaaaaaacctltg caagcctaalya      | 28530 |
| Dd | 3394 | TTATAGAGGAAGATCAGAGCGCAAGAGACAGACACTGAAGAAAACTTGTGCMAAGCCTAATGA    | 34533 |
| Oy | 2851 | aaecaaaacttacttlttgaagtlgcacacataatylgcacccactlaagaagtlsgtltya     | 29130 |
| Dd | 3454 | AACCAAACTACTTTTGGAAAGTCGACACATATATATGAGCCCACTTAAGATGAGTTTGA        | 35133 |
| Oy | 2911 | ctgcaaaagccctgggactaatctctcaba bgtlbaccctggaaaaaga tyrgactaagcct   | 29730 |
| Dd | 3514 | CTGCAAAAGCCTGGGCTTATTTCTCTATGTTGACTGGAAGAAAGTGTGCCACTAGGECT        | 35733 |
| Oy | 2971 | gatttgagccctctcygctcbyccacactaa cacaactyaaccctlygtclayggagacaagt   | 30330 |
| Dd | 3574 | GATTGAGCCCTTCTGTGCTGCACACTAAACACAGAAACCTGTCTATGSGAGACAAAGT         | 36333 |
| Oy | 3031 | gaagagtaagaagatttgtctgtltttcaacalc ttltga agaccaaac gtygtact       | 30930 |
| Dd | 3634 | GACAGTACAGGAATTTGTCTGTTTTCAACCATTTTGTGTGAGACCAAAACCTGCTACTT        | 36933 |
| Oy | 3091 | caetgaaataatatygaagaaacetycagagctccctbgcaatlcacagatlyggaagatccac   | 31530 |
| Dd | 3694 | CACGTGAATAATVTGGAAGAAGAAACGTGAGGGCTCCCTGTGCATATCCAGATGGAAGATCCAC   | 37533 |
| Oy | 3151 | ttttaagagaatlatcogcttccatqycaa tcoaatyygctatacabaatylgalactactcbyg | 32130 |

[illegible]

QY 4291 acctgtgtgaactctctagaccacccgcttactgactcgtactcgtactcgaatcaccacca 4350  
|||||  
Db 4894 ACCTGTGTGAACCTCTTAGACCAACCGTTACTGACTGCTACCTCGAATTCACCCCA 4953  
QY 4351 gaattgggtgcacgaatgtccctgaagagatgtgaggtctcgtggtcgtcgaagcacaagacct 4410  
|||||  
Db 4954 GAGTTGGGTGCACCAAGATTGCCCTGAGATGAGAGTTCGGCTGCGAGGACAGGACCT 5013  
QY 4411 ctactga 4417  
|||||  
Db 5014 CTACTGA 5020

RESULT 8  
LOCUS AR071306 5035 bp DNA PAT 18-FEB-2000  
DEFINITION Sequence 1 from patent US 5910481.  
ACCESSION AR071306  
VERSION AR071306.1 GI:7222194  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5035)  
AUTHORS Voorberg,J.J.  
TITLE Hybrid proteins with modified activity  
JOURNAL Patent: US 5910481-A 1 08-JUN-1999;  
FEATURES  
Source 1..5035  
Location/Qualifiers

BASE COUNT 1484 a 1127 c 1110 g 1314 t  
ORIGIN

Query Match 84.8%; Score 3743.8; DB 9; Length 5035;  
Best Local Similarity 87.5%; Pred. No. 0;  
Matches 4365; Conservative 0; Mismatches 52; Indels 570; Gaps 2;

QY 1 catcaaatagagctctcgaactgtctctctctgtgaccttggcgatctgctttagtg 60  
|||||  
Db 34 CATGGAATAGAGCTCTCACCCTGCTTCTTCTGTGCTTGGCTTTGGCATCTGCTTAATGTC 93  
QY 61 caccagaatactactcctggtgctcagtggaactgtcatalgtggaactatgtcaaatgtatct 120  
|||||  
Db 94 CACCAAGAATACATCTACCTGGGTGAGTGAATGCACTGATGAGATGATATGCAAAAGTATCT 153  
QY 121 cgtgtgagctgctgtgtgacgcaagaattctctctctagatggtgccaaaatcttccatcaa 180  
|||||  
Db 154 CGGTGAGCTGCTGTGACGCAAGATTTCTCTAGAGTGCAGAAATCTTTCCATTTCAA 213  
QY 181 caactcaatgtgtatacaaaaagactctgtttagaatcaggttcaactcttcaaat 240  
|||||  
Db 214 CACCTCACTGCTGTATCAAAAAGACTCTGTTTGTAGAAATTCACGGATCACCTTTTCAACAT 273  
QY 241 cgttaagcacaagccacccctgtagtgcgtcgtcgtacgtccatccacatccaggtcgaagtta 300  
|||||  
Db 274 CGTTAAGCCAAAGGCGACCCCTGGATGGGTCTGCTAGTCTTACATCCAGGCTGAGGTTTA 333  
QY 301 tgaatacagtgatcattacacttaagaacatggtcttccatctcgttcaagtcttcaagtct 360  
|||||  
Db 334 TGATACAGTGTCTCATTTACATTAAAGACATGCTTCCATCTCTGTCAGTCTTCATAGCTCT 393  
QY 361 tgggtatccctactggaagaactctggaaggaactgaatatgataatgaatgaaccgaactcaag 420  
|||||  
Db 394 TGGGTATCTCTACTGTGAAGGCTTCTGAGGGAGCTGAATATGATGATCAACCAAGTCATAAG 453  
QY 421 ggaagaagaagaatgaataaagcttccctggttgaagcacaatacatatgtctggaagttcc 480  
|||||  
Db 454 GGAGAAAGAAGATGATTAAGTCTTCCCTGTTGGAAGCCATACATATGTCGAGAGTCTCT 513  
QY 481 gaaagaagaatgtccaatgtgacctctgaaccactgtgaccttactcctcatccttctca 540  
|||||

Db 514 GAAAGAAATGTTCCAAATGCGCTCTGACCCACTGTGCTTACCTACATATCTTTCTCA 573  
QY 541 tbtgaccctgtgtaaagaacttgaaatcaaggtctcatgtgaagccctactagatgtagaaga 600  
|||||  
Db 574 TGTGAGACTGTGTAAGAACTTGAATTCAGGCTCTCATTTGGAGCCCTTACTAGTATGAGGA 633  
QY 601 aggaagctctgccaagaagaaacacagaaccttgacaatattactactcttctgtct 660  
|||||  
Db 634 AGGAGTCTGCGCAAGGAAGAACACAGACCTTGACAAATTTATTAATCTTTTGTCTGT 693  
QY 661 atttgatgaagggaaaagtgtgcatctcgaagaacaaagaactctcttgatgacagatagaaga 720  
|||||  
Db 694 ATTTGATGAAGGAAAGAAAGTTGGCACTCAGAAACAAAGAACTCTTGATGACGAGATAGGGA 753  
QY 721 tgcgtcatctgctcgagccctgagcctaagaatgacacagtcacatgtgttattgtacaacagctc 780  
|||||  
Db 754 TGCTGATCTGCTCGGCGCTTGCCCTTAATATGACACAGTCAATGTTATGTAACAGGTC 813  
QY 781 tctgccaagctctgatatgtgatagtccacagaagaatcagtcatalgtgcatgtgatatgtgag 840  
|||||  
Db 814 TCTGCCAGGTCTGATTTGATGATGTCACAGGAATCAGTCTATTGGCATGTGATTTGGAATGG 873  
QY 841 caacactcctgaagtgtcactcaatatctcgtgaaggtcacaacattctgtgtgaagaacca 900  
|||||  
Db 874 CACCACTCTGAAGTGCACATCAATATCTCTGAAAGTGCACACATTTCTGTGAGGAACCA 933  
QY 901 tgcgaagcgtctcttggaaatctgcgcaataactctctctgctgctcgaacaactctgat 960  
|||||  
Db 934 TCGCCAGCGCTCTTGGAAATCTCCGCATTAATCTTCTTACTGCTCAACACCTTTGAT 993  
QY 961 ggaacttgagacagtttctactgttctgtcatalctctctccacaacaatgtatgtgcatgtga 1020  
|||||  
Db 994 GGACCTTGACAGTTTCTATCTGTTTGTCAATATCTCTCCACCAACATGATGAGATGGA 1053  
QY 1021 agcttatgtccaagttagaagaagctgtccagaggaaccccaactcagaatgaaataatgtga 1080  
|||||  
Db 1054 ACCTTATGTCAAAAGTAGAGAGCTGTCCAGAGGAACCCCACTACCAAGAAATAAATGA 1113  
QY 1081 aagaagcgaagaactatgatatgatacttactgactcgtgaatgtgaatgtgtgtcaggttga 1140  
|||||  
Db 1114 AGAAGCGAAGACTATGATGATGATCTTAATCTTAATCTGAATGATGTGTGACAGTTTGA 1173  
QY 1141 tgaatgaactctcctctcttaccatccaaattcgtcagltgccaagaagcatcctaaac 1200  
|||||  
Db 1174 TGATGACACTCTCTCTCTTATCCAAATTCGCTCACTTGCCAGAGACATCTCTAAAC 1233  
QY 1201 ttgggttaacttcatctgtcgtcgtgaagaggaagactgggaactatgctcccttaagctcctgc 1260  
|||||  
Db 1234 TTGGGTACATTTACATTTGCTGTGAAGAGGAGACTGGGACTATGCTCCCTTAGTCTTGGC 1293  
QY 1261 ccccgatgacagaagtataaagaatcaatatlttgacaagtgtgccccctcagcggaattgttag 1320  
|||||  
Db 1294 CCCCGATGACAAAGTTATTAAGAAAGTCAATATTTGAACAATGGCCCTTCAGCGGATTTGTAG 1353  
QY 1321 gaagtacaataaagctcgaattatgtgcatlacagatgaacacattlaagactcgtlbaagc 1380  
|||||  
Db 1354 GAAGTACAAAAGGTCGCGATTTATGATGACATGAAACCTTTAAGACTCGGAAGC 1413  
QY 1381 tatcaagcatgaatcaagaatcttgggaccttacttatagtgggaagtgtgaagacaact 1440  
|||||  
Db 1414 TATTCAGCATGTAATCAGGAATCTTGGGACCTTTATGTTATGGGAAAGTTGAGAGACACT 1473  
QY 1441 gttgattatatttaagaatcaagaagaagacatataatactacccctcaggaataac 1500  
|||||  
Db 1474 GTTGAATATATTTAAGAAATCAAGCAAGACATATATACCTTACCTTCAGGAATTCAC 1533  
QY 1501 tgaatgccgtctctgttattcaaggagattacccaagaagtgtlaaacaatttgaagaattc 1550  
|||||  
Db 1534 TGATGTCGCTCTTGTATTCAGAGAGATTTACCAAAAGCTGTAAACATTTGAAGGATTT 1593  
QY 1561 tccaattctgccaaggaagaatatttcaaatataatgtgacatgtactgtgaagaagtgtggcc 1620  
|||||  
Db 1594 TCCAAATTTGCGCAGGAAGAAATTTTCAATATTAATGAGACAGTGTAGAAAGATGGGCC 1653

|    |      |  |      |
|----|------|--|------|
| Qy | 1621 | aactaaatcaatcctcggtgcctgacccgcctattactagtgttgtaatatggagag      | 1680 |
| Db | 1654 | AACATAATCAGATCTCTGGTGCCCTGACCCGCTATTACTCTAGTTTCTGTTAATATGGAGAG | 1713 |
| Qy | 1681 | agatctagctcaggaactcattgcccctctccatcctatcgtctacaaagaatcgtatga   | 1740 |
| Db | 1714 | AGATCTAGCTTCAGGACTCATATGGCCCTCTCTCATCTGCTGACAAAGATCTGTAGATCA   | 1773 |
| Qy | 1741 | aagaggaaccagataatgtcagaacaaagggaatgtcatccctgtttctgtatttga      | 1800 |
| Db | 1774 | AAGGGAACCCAGATATGTCAGACAAAGAGAAATGATCATCTGTTTCTGTGATTTGATGA    | 1833 |
| Qy | 1801 | gaaccgaagctggtgacctcagaagaatatcaaacgctttctcccaatccagctgagat    | 1860 |
| Db | 1834 | GAACCGAAGCTGTAACCTCAGAGAAATATACAGCCTTTCTCCCAATCCAGCTGGAGT      | 1893 |
| Qy | 1861 | gcagcttgagatccagaagttccaaagcctccaacatcatgacagacatcaatggtatgt   | 1920 |
| Db | 1894 | GCAGCTTGAGGATCCAGATGCCAAGCCTCCAAACATCATGACAGCATATGGCTATGT      | 1953 |
| Qy | 1921 | tttgatagtttgcaagttgtcagttgttctgcatgaggtgagcatcctggtatcctaa     | 1980 |
| Db | 1954 | TTTTGATAGTTTGCAAGTTGTCAGTTGTTGCATGAGGTGGCATATGTCATATTTCTAAG    | 2013 |
| Qy | 1981 | catlgagacacagactgacttcccttctgtcttcttctctctggaatataccttcaaa     | 2040 |
| Db | 2014 | CATTGGAGCACAGACTACTCTCTTTCTGTCTTCTTCTCTGATATACCTTCAACACAA      | 2073 |
| Qy | 2041 | aatgctatagaagacacactcaacctatcccatctccagaagaactgtctcatgctc      | 2100 |
| Db | 2074 | AATGCTTATGAAGACACACTCACCTTATTCCTCAAGGAAACTGCTTCAATGTC          | 2133 |
| Qy | 2101 | gatggaacacccaggtctatgagatctcggggtgcacaaactcagaacttcgggaacag    | 2160 |
| Db | 2134 | GATGGAACCCAGGCTATGAGATTCGGGGTGCACAACTGCAACATTTGGAACAGAG        | 2193 |
| Qy | 2161 | catgacgccttactgaagtttctagttgtgacaagaacagtgatattatcacggga       | 2220 |
| Db | 2194 | CATGACGCCCTTACTGAGAGTTTCTAGTTGATTCAGAGGGGAGAGGCGCGACTA         | 2253 |
| Qy | 2221 | caatgaatgaatattca-----gcatactgtcgtgataaaacaatgac               | 2265 |
| Db | 2254 | TCTGGAACCTGGAGAAATATCACTGTAAGAGAGACTACATGACATGCGACACTCT        | 2313 |
| Qy | 2266 | catggaacacaaaggtctccccaagaattcaagaacccctagaactgagcaaaagaatt    | 2325 |
| Db | 2314 | GATTGAACCAAGAGCTTCTCCAGAAATCAAGACACCCCTAGCACTAGGCAAAAGCAATT    | 2373 |
| Qy | 2326 | taatgacac-----   | 2334 |
| Db | 2374 | TAAATGCCACCAATTCACAGAAATGACATAGAGAACGTGACCTTGTGGCACACAG        | 2433 |
| Qy | 2335 | -----  | 2334 |
| Db | 2434 | AACACCTATGCTAATAATACAAAATGTCCTCTAGTATTTGTTGATGCTTTGGCACA       | 2493 |
| Qy | 2435 | -----  | 2334 |
| Db | 2494 | GAGTCTACTCACANGGGCTATCTTATCTGTGATCTCCAGAAAGCCAAATATGAGACTTT    | 2553 |
| Qy | 2535 | -----  | 2334 |
| Db | 2554 | TTTCTGATGATCCATCACTTGAGAGCAATAGACAGTAATTAACAGCTGTCTGAATGACACA  | 2613 |
| Qy | 2614 | CTTCAGGACACAGCTCCATCAGATGGGACATGATATTACCCCTGATGAGGCTTCCA       | 2673 |
| Qy | 2635 | -----  | 2334 |
| Db | 2674 | ATTAAGATTAATGAGAAACTGGGACAACTGCAGATCTCTTGTGGATTAACCACTA        | 2733 |

---

|    |      |  |      |
|----|------|--|------|
| Qy | 2335 | -----  | 2334 |
| Db | 2734 | TGTTACTCAGATACCAAAAGAGAGTGGAATCCCAAGAGAAAGTACACAGAAAAAACACC    | 2793 |
| Qy | 2335 | -----  | 2334 |
| Db | 2794 | TTTTAAGAAAAGGATACCATTTTGTCCCTGAACGCTTGTGAAAGCAATCATGCATAGC     | 2853 |
| Qy | 2335 | -----  | 2334 |
| Db | 2854 | AGCAATTAATGAGGACAAAATAAGCCGAAATAGAAATGACCTGGCCAAAGAGGTAG       | 2913 |
| Qy | 2335 | -----  | 2334 |
| Db | 2914 | GACTGAAGGCTGTGCTCTCAAAAACCAAGTCTTAAAGCCCATATACAGGAAATATAC      | 2973 |
| Qy | 2971 | tcgtactactcttcagtcagatcaagaagaatgtactatgatatccalatcagttga      | 2430 |
| Db | 2974 | TGTTACTACTCTTCAGTCAGATCAAGAGAAATGACTATGATGATACCATATCAGTTGA     | 3033 |
| Qy | 2431 | aatgaaagaagaagatttgacattatgatatgagatgaaatcagaagcccccgagctc     | 2490 |
| Db | 3034 | AATGAAGAAGGAAGATTTTACATTTATGATGAGGATGAAATACAGAGCCCGCACCTT      | 3093 |
| Qy | 2491 | tcaaaagaacacagacactatttactgtgcagtgagagagctctgagattagagat       | 2550 |
| Db | 3094 | TCAAAAAGAAAACAGACACTATTTATTTGCTGCAGTGGAGAGGCTCTGGGATTTGGGAT    | 3153 |
| Qy | 2551 | gagtagtccccacatgttctaaagaaacaggtcacaagtgagtgctccctcagttcaa     | 2610 |
| Db | 3154 | GAGTAGTCCCCACATGTTCTTAAGAAACAGGGCTCAGAGTGGGAGTGCTCCCTCAGTTCAA  | 3213 |
| Qy | 2611 | gaaagttgttccagaagaatttactgagtgcctccttactcagcccttatacgttgaga    | 2670 |
| Db | 3214 | GAAAGTTGTTTCCAGGAATTTACTGATGAGCTCCCTTACTCAGGCCCTTAAACCTGGAGA   | 3273 |
| Qy | 2671 | actaaatgaacatttggagactccctggggccatatataagaagcagaagtgtgaagataat | 2730 |
| Db | 3274 | ACTAATGAACATTTTGGGACTCCTGGGGCCATATATAGACAGCACTTAAGATTAATAT     | 3333 |
| Qy | 2731 | catgttacttccagaatcagaagcctcgtccctatcttctctatctagcttattc        | 2790 |
| Db | 3334 | CATGTTACTTTTGAATATCAGGCGCTCTGCTTATCTCTTATCTGACCTTATTTTC        | 3393 |
| Qy | 2791 | ttatgagaagaatccagaagcagaagcagaactagaaaaaacttgtcaagcetaatga     | 2850 |
| Db | 3394 | TTATGAGGAAGATTCAGAGGAGAGGAGCAAGCACTAGAAAAAATCTTGTCAAGCCTAATGA  | 3453 |
| Qy | 2851 | aaccaaaacttacttttggaaagtgcacatcatatgagccccaactaaagatgattga     | 2910 |
| Db | 3454 | AACCAAAACTTACTTTTGGAAAGTGCAAACATCATATGGCACCCACTAAGATGAGTTGA    | 3513 |
| Qy | 2911 | ctgcaagccttgggtctatttctctgtatgttgcacttggaaaaaagatgtgactcaggct  | 2970 |
| Db | 3514 | CTGCAAGCCTGGGCTTATTTCTCTGATGTGACCTGAAAAAAGATGTGACACTAGGCGCT    | 3573 |
| Qy | 2971 | gatttgacccctctgtgtcgcacactaaacaacttgaaccccgctcatatgtggagacaagt | 3030 |
| Db | 3574 | GATTGAGCCCTTCTGTGTCGCCACACTTAACACTGAACCTGTGCTATGTGGAGACAAAT    | 3633 |
| Qy | 3031 | gaacgtacaggaatttgcctgttttccacatcttctgtatgagaccaaagcgtgtactc    | 3090 |
| Db | 3634 | GACAGTACAGGAATTTGCTGTTTTTACACATCTTTGATGAGACCAAAAGCTGGTACTT     | 3693 |
| Qy | 3091 | cactgaaatatalgaaagaacatgcaaggtcccttgcataatccagatggaagatccac    | 3150 |
| Db | 3694 | CACGTGAATAATATGAAAGAAATGCAAGGCTCCCTGCAATATCAGATGGAAGATCCAC     | 3753 |
| Qy | 3151 | ttttaaagaataatctgccttcatatgcaatcaatggttataatgatatcactactg      | 3210 |
| Db | 3754 | TTTTAAAGAAATTAATTCCTTCATGCAATATGCGTATCATATGATGATACACTTCTGG     | 3813 |
| Qy | 3211 | cttagtaatgtccagatcaagaagatcgatgtatctcgtacgaatgagcagcaatga      | 3270 |

|||||  
Db 3814 CTATGATATGCTCAGATGCAAGATTCGATGATCTGCTCACACATGGGAGCAATGCA 3873  
3271 aaacatcatctatctatcttcacgtgacatgctgctacgtacgaaaaaagagagta 3330  
Db 3874 AAATCATCTATTCTTCTTTCAGATGACATGCTGCTACGTGACAAAAAAGAGAGTA 3933  
3331 taataatgcaactgtacaaatctctacccagtgcttcttgagacagtgaatgtaccatc 3390  
Db 3934 TAAATATGGACTGTACATCTCTATCCAGGTGTTTTCAGACAGTGAATGTATACATC 3993  
3391 caaagctggaaatttgagggtggaatgccttatgvcgagcaatctacatgctggatag 3450  
Db 3994 CAAAGCTGGAAATTTGGCGGTGGAATGCTTATGGCAGACATTCACATGCTGGATGAG 4053  
3451 cacacctttcttggtgacagcaatgaatgtcacagctccctggagaaatggtcttgaca 3510  
Db 4054 CACACTTTTCTGTGTGACACAAATAGTGTACACTCCCTGGGAATGCTTGTGACAA 4113  
3511 catlagagatttccagattacagcttcagacaatatgacagtgggcccacaagctgac 3570  
Db 4114 CATTAGAGATTTTTCAGATTACAGCTTCAGACAATATGAGACGTGGGCCCAAACTGGC 4173  
3571 cagaactcatatctccggatcaatcaatgctgtgagacccaaggagcccttcttgat 3630  
Db 4174 CAGACTTCATTAATTCGGATCAATCAATGCTGAGACCAAGGAGCCCTTTTCTTGAT 4233  
3631 caaagtgaatctgtgtgacccaatgatatctacagagcaatcaagaccagggtgcccgtca 3690  
Db 4234 CAAAGTGGATCTGTTGGCACCATAATGATTAATTCACGGCATCAAGACCCAGGCTGCCGTCA 4293  
3691 gaagttctccagcctctacatctctcagttctatcaatcagttatgcttgatgagagaa 3750  
Db 4294 GAAATGTCACGACCTCTCATCTCTCAGTTTATCATCATGTACTGTTGATGGAGAA 4353  
3751 gtggcagaactatcagaggaatctccactggaaccttaatgctctcttgcaatgtga 3810  
Db 4354 GTGGCAACTATATGAGGAAATTCACACTGGAACCTTAATGCTTGTGGCAATGTGA 4413  
3811 ttcatctggagataaacaacatatcttaaccctccaatatgtctgcatataccgctt 3870  
Db 4414 TTCTATCTGGGATTAACCAATATTTTAACTTCAATTTATGCTCGATACATCCGTTT 4473  
3871 gcaaccaactcatatagatctgcagacactcttcgcacatggagtgtgagctgtgact 3930  
Db 4474 GCACCAACTCATTAATAGCATTCGAGCAGCTCTTGATGAGTGTGATGGCTGTGATTT 4533  
3931 aaatagttgcagcatgcatctgggaaatgagaaagtaaacataatcagatgcagatcac 3990  
Db 4534 AAATAGTTGCGATGCGCATTTGGGAAATGAGAAATCAATATCATGATGCAACATTAAC 4593  
3991 tgcctcatcctactcttaacaatatgttgcacacctggtctcccttaaaagctcgactca 4050  
Db 4594 TGCCTTACCTCACTTATACCAATATGTTTGGCACCCTGGTCTCCTTAAAAAGCTCGACTTCA 4653  
4051 cctccaaagggaagaaatgacctggaacctcaagtgtaataatccaaaaagatggctga 4110  
Db 4654 CCTCCAGAGGAGGAGATATGCTGAGACCTGAGACTGATTAATCCAAAAGATGGCTGCA 4713  
4111 agtggacttccagaagaacatgaaagtacacagggtaactactcaaggaagtaaaatctct 4170  
Db 4714 AGTGGACTTCCAGAAAGCAATGAAGATCAAGAGTAATCTACTCGGAGAGTAATATCTCT 4773  
4171 gcttaccaagcatgtatgtgaaggatctccatctccagcagtaagaatgacacagtg 4230  
Db 4774 GCTTACCAAGCATGTATGTGAAGAGACTTCTCATCTCCAGCACTCAAGATGGCCATCAGTG 4833  
4231 gactctcttttccaagaatggcaagtaaggttttccaagggaatccaagactcttcac 4290  
Db 4834 GACTCTCTTTTTCAGATGCGCAAGTAAAGTTTTCAGGGAATCAAGACTCTCTTAC 4893  
4291 aactgtgtgaactctctagaccacgcttaactgactgactaacttgaaatcacccca 4350  
|||||

Db 4894 ACCTGTGTGAACCTCTAGACCCACCGTTACTGACTGCTACCTTGGAATTCACCCCA 4953  
4351 gattggtgtgaccagatctccctgaagatgaggtcttggtgctgcagacagacct 4410  
Db 4954 GAGTTGGGTGCACCAAGATTGGCTGTAGATGAGAGTTCTGGGCTGGAGGACAGACT 5013  
4411 ctactga 4417  
Db 5014 CTACTGA 5020  
RESULT 9  
AR029098 4334 bp DNA PAT 29-SEP-1999  
LOCUS AR029098 Sequence 38 from patent US 5859204.  
DEFINITION AR029098  
ACCESSION AR029098.1 GI:5941071  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4334)  
AUTHORS Lollar, J. S.  
TITLE Modified factor VIII  
JOURNAL Patent: US 5859204-A 38 12-JAN-1999;  
FEATURES Location/Qualifiers  
source 1..4334  
BASE COUNT 1109 a 1126 c 1085 g 1014 t  
ORIGIN  
Query Match 72.2%; Score 3189.8; DB 9; Length 4334;  
Best Local Similarity 83.5%; Pred. No. 0;  
Matches 3692; Conservative 0; Mismatches 637; Indels 90; Gaps 3;  
3  
2 atgcaaatagagctctccacgtgcttctctgtgcttcttgagatctgcttagtgc 61  
Db 3 ATGCAAGTACAGCTCTCCACCTGTGTCTGTGTCTCTGTCCTTGCACATCGGCTTAGTGC 62  
62 accagaatatctactctgtgtgctgagtggaactgcatgagagctatagcaaaagta---t 118  
Db 63 ATCAGAGATACCTACTGCGCGACGTGAGACTGTCTGGACTACCGGCAAGATGAATCTC 122  
119 ctctgtgagctgctgtgtgacgcaagattctctcctgaagtgccaaatcttccaatc 178  
Db 123 CTCCTGAGCTGACAGTGTGACACACGATTTCTGCTACAGGCGCAGAGGCTTCGCTTG 182  
179 aacacctagtcgtgtacaaaagaactctgtgtgtagaattcacggttcaactttcaac 238  
Db 183 GCGCCGTGACGTCCTGTACAAAAGACTGTGTGTGAGAGTTCACGAGATCAACTTTTCAGC 242  
239 atcgtctaaagcaaggccacctgagtggtgtgtgtagtctaccatccagcagctgaagt 298  
Db 243 GTTGGCAGGCCCGAGGCCACATGATGGGTGCTGTGGTCTTACCATTCAGAGCTGAGGTT 302  
299 tatgaataagtgatcaacttaagaacatggcttcccatcctgtcagctctcatgct 358  
Db 303 TACGACAGGCTGTGTTACCTGTGAAGAACATGGCTTCTCATCCGTTAGTCTTCAAGCT 362  
359 gtgtgtatctactatgaaagctctgagagagctgaataatgatatgacagacgtcaa 418  
Db 363 GTGGCGTCTCTCTTGTGAATCTTCCAAAGGCGCTGATGTGAGATGACACACAGCCAA 422  
419 agggagaaagaagatgataaagcttctccctgtgtggaagcaataatgtctgagcagt 478  
Db 423 AGGGAAGAGAAAGAGATGAAGTCTTCCGGTAAAGCCAAACCTACGCTGTGCAAGTTC 482  
479 ctgaaagagatgtgtccaatgacctcagccactgtgcttaccatctactatctctc 538  
Db 483 CTGAAGAGAAATGTGTCCACACAGCTGTACCCACATGTCTCACCTTACATACCTGTCT 542  
539 catgtgacctgtgataaagaacttgaaatcagagctcatgtgagccctactagatgtaga 598  
|||||

|    |      |   |      |
|----|------|---|------|
| Db | 543  | CACGTGAACTGCTGAAAGACCTGAAATTGCGGCTTCATTGGACCCCTGCTGTTTGTAGA     | 602  |
| Oy | 599  | gaaggaagtcctgagcaagaaagaaacagacccttgacaaattalatactacttctgct     | 658  |
| Db | 603  | GAAGGAGTGTGACACAGAAAGACCAGAACCTGACAGATTTGTACTCTTTTCT            | 662  |
| Oy | 659  | gtattgatgaaggaaagtctgagacatcaagaacaagaactccttgatgcagatagg       | 718  |
| Db | 663  | GCTTTTGATGAAGGAAAGTAATGTGGCATCTGACAGAAATAGATCCTTGACACGGGCTAG    | 722  |
| Oy | 719  | gaigctgacatctgctcgggctcgtgacctaaagagcacagtgcaatggttatgtaacag    | 778  |
| Db | 723  | GATCCCGACCTGCGACAGGGCCAGCCTGCAATGACACAGTAAAGGCTATGTCAACAGG      | 782  |
| Oy | 779  | tctctgcagatcgtgatgtatgcacaaagaaatcagatctatctgacatgatatggaat     | 838  |
| Db | 783  | TCTCTGCGAGGTCTGATCGSATTGCTCAATAGAAATCAGTACTGGCAGCTGATTGGAATG    | 842  |
| Oy | 839  | ggacacacccctggaagtgcatcacaatalctcctcgaaagtgcaacacattcttgtagaac  | 898  |
| Db | 843  | GGCACACGAGCCGGAGAGTGCACCTCATTTTCTTTGAAGGCGACACGTTTCTGTGAGCAC    | 902  |
| Oy | 899  | catcgcaagcgctccttggaatctctgcacaataactttcccttaactgctccaacctctg   | 958  |
| Db | 903  | CATGCGCAGGCTTCCTTGGAGATCTGCGCACATTAACCTTCTCATCGTCAGACATTCCTG    | 962  |
| Oy | 959  | atgagacctggaagcttctactgcttcttgatctctccacacaatgatgagatg          | 1018 |
| Db | 963  | ATGGAACCTTGCCAGTGTCTCTACTGTTTGTCAATCTCTCCACACCACTATGTGGCATG     | 1022 |
| Oy | 1019 | gaagcttatgtlcaaaagtagaagcagctgtccagaaggaaccccaactacagatgaaataat | 1078 |
| Db | 1023 | GAGGCTCAGCTCAGATGATGAAGAACTGCGCGAGAGAGGCCCACTGCGAGGAAGCTGAT     | 1082 |
| Oy | 1079 | gaagaagcgggaagactatgtagtacttactgattctgaaatgagatgtagtgaagtt      | 1138 |
| Db | 1083 | GAAAGA---GGAAGTTATGATGATGACAAATTTGTACGACTCGACATGAGAGTGTGCGGCTC  | 1139 |
| Oy | 1139 | gaibaibaacaacctctcctcttaccacaattgctcagtgatggccaanaaagatctcaaa   | 1198 |
| Db | 1140 | GATGCTACACAGCTGTCCTCTTATTCACAAATCCGCTGCTGTTGCCAABAAGATCTCCAAA   | 1199 |
| Oy | 1199 | acttggttacattlactatgtcgtctgtaagagaggaactggaactatgctccctagtcctc  | 1258 |
| Db | 1200 | ACCTGCTGACACTAATCTCTGCAGAGGAGAGAGACTGAGACTAGCCGCCGCGGTCCC       | 1259 |
| Oy | 1259 | gcgcccgatgacagaaagtataaaagtcacatattgaaacatgcccctcaagcgatgtgt    | 1318 |
| Db | 1260 | AGCCCCAGTACAGAAAGTTATAAAAGTCTCTACTTGAAACAGTGGTCTCACGSAATTTGT    | 1319 |
| Oy | 1319 | aggaagaagaacaaaagtcogattatgtagcatcacagaatgaaacctttaagacctcgtaa  | 1378 |
| Db | 1320 | AGGAAATACAAAAGCTGCATTCGTGCTTACACGGATGTAAACATTTAAAGACTGCTGAA     | 1379 |
| Oy | 1379 | gcatltaagaatgaatcaagaaatcttgggaaccttacttlatggygaagtctgagacaca   | 1438 |
| Db | 1380 | GCTATTCGGATGAATCAGGAATCCGCGGACCTTATCTTATGAGAAGTTGGACACACA       | 1439 |
| Oy | 1439 | ctgttgatlatltaagaatcaagaagcaagacacatabaacatctacccctcaagyaatc    | 1498 |
| Db | 1440 | CTTTTGATTTATTTTAAGATTAAGGACGCCACCATTAATACATCTACCTCATAGGAATC     | 1499 |
| Oy | 1499 | actatgtccgtctcttctgtatctcaaggaatataccaanaagtgtaaacatcttgaagat   | 1558 |
| Db | 1500 | ACTATATTCAGCCCTTTGGCACCCAGGGAACCTTTAAAGGTTTGAAACATTTTGAAAC      | 1559 |
| Oy | 1559 | tttcaaatctgcgaaggaagaatcatcaataataatgagacagttaactgtagaagatg     | 1618 |
| Db | 1560 | ATGCGCATTTCTGCCAGAGAGACTTTCAAGTATTAATGACACGTACTCTGTGGAAGATGG    | 1619 |
| Oy | 1619 | ccaactaatcagatccctcggtcctgacccgctatlaactagtttctgtlaatatgag      | 1678 |

|   |   |      |  |      |
|---|---|------|--|------|
| D | b | 1620 | CCACCAAGTCCGATCCTCGGTGCCTGACCCTACTACTCGACGCTCATTTAATTAAAG            | 1679 |
| O | y | 1679 | agagatcgaacttcaaggacatcatctggccctctcctcatctgtgtacaagaactgtagat       | 1738 |
| D | b | 1680 | AAAGATCTGGCTTGSGGACATTAATGGGCCCTCTCATCTGTGGTACAAGAATCTGTAAAC         | 1739 |
| O | y | 1739 | caaagaagaaaccagataatgltcacagaagaagaatgltcatccgtlcttgtaattgat         | 1788 |
| D | b | 1740 | CAAAAGAGCAAAACGATGATGTACAGACAAGAGAAAGTCATCTGTTTTCTGTAATGCAT          | 1799 |
| O | y | 1799 | gagaaacgaagctgtgtatcccacaagaagaatatcaacgctttctccaatccagctgt          | 1858 |
| D | b | 1800 | GAGATTCAAACCTGTGTAACCTCGACGAAGAAATTAATCACGCTCTCTCCCAATCCGATGGA       | 1859 |
| O | y | 1859 | gtcgagcttgtgaga tccaagattccaagccctccaatcatgtcacagcatcaatgtgcata      | 1918 |
| D | b | 1860 | TTCACGCCCCGAGATCCAAGATTCCAAAGCTTCAATCATGTACACAGCATCAATGGCTAT         | 1919 |
| O | y | 1919 | gttlttgatagttttcagattgttgcagtlltgtlgtcaltyagtyggcaltygtatactcta      | 1978 |
| D | b | 1920 | GTTTTGTATACCTTCGACGCTGCTGGTTGTTGTTSCACAGAGTGGCAATCGTAACATTTCTA       | 1979 |
| O | y | 1979 | agcaattgagacagacagactgacctcttcctgtcttcttctctcgtatataccctcaaac        | 2038 |
| D | b | 1980 | AGTGTGGAGACAGACGGAGCTTCTCTCGCTTCTTCTCTGGCTACACCTTCAAAAC              | 2039 |
| O | y | 2039 | aaaatgtgctctatgagaacacactacccttatccaattccaaggagaactgttctctg          | 2098 |
| D | b | 2040 | AAAAATGGTCTATGAAAGCACACTCACCCCTYTCCCTTTCAGAGAAAAGSTCTTTCATG          | 2099 |
| O | y | 2099 | tcgatggaaaaacccccagttctctatgtatcttgygtgccaacactcagacttctgagaaga      | 2158 |
| D | b | 2100 | TCATAVGAAAAACCCAGGTCTCTGGTGTCTTAGSTGGCCAACATCGACTGGGAAACGA           | 2159 |
| O | y | 2159 | ggcatgaccgcccttactcgtaaagtttctagtgtgtacagaagaacactggtgatattaagag     | 2218 |
| D | b | 2160 | GGGATGACAGGCTTACTGTAAGGTGTATAGTTGTGACAGGAGCATTTGGTATTAATTATAC        | 2219 |
| O | y | 2219 | gacagttatgaagatatttcaagaacttgcgtgagtaaaacaatgtgccaattgaaccaaga       | 2278 |
| D | b | 2220 | AACACTTATGAAGTATTCCAGGCTTCTTGCTGATGGAAGAAATGTCTTTAAACCCAGA           | 2279 |
| O | y | 2279 | agctctccccaagaatlcaagaacacccctagcactlagcaaaagcaatltaatgtccaccca      | 2338 |
| D | b | 2280 | -----  | 2279 |
| O | y | 2339 | ccagatcttgaaacgcatacaacgggaataactcgttactactctcgaigtacagataagag       | 2388 |
| D | b | 2280 | -----GACATAAGGCTTCTTACCTTTCACGCGGAGGAAAGAC                           | 2315 |
| O | y | 2399 | gaattactatatagttaccataatgtttaagaatbaagaagaagaatttgaacattat           | 2458 |
| D | b | 2316 | AAAAATGGCTATGTGATTAATCTTCTCACTACGAAACAGAAAGGAAAGATTTTACATTTAC        | 2375 |
| O | y | 2459 | galtaagatgaanaalctcaagaccccccgagcttccaagaagaanaacacgacatattlat       | 2518 |
| D | b | 2376 | GGTAGAGATGAATAATTCAGACACCTCGAGGTTTCAABAAGAAACCGACACGATTTCAAT         | 2435 |
| O | y | 2519 | gctgctcagttgagaagcctctvggattatggatgtagtlaagctccccaatgttctaaagaac     | 2578 |
| D | b | 2436 | GCTCGCTTGAGCGAGCTCTGGGATTTACGGGATGAGCGAATCCCCCGGGCGCTAAGAAAC         | 2495 |
| O | y | 2579 | agggctgaagatgtagcagtgctcccttaagttcaagaagaatgtgttcttcagaagaatttactgat | 2638 |
| D | b | 2496 | AGGGCTCAGAACGGAGAGGTGCTCTCGGTTCAAAAGTGTCTTCGGGAAATTTGCTTAC           | 2555 |
| O | y | 2639 | ggccccttactcaagcccttataacgttgtagaactaaatgaaactttggaotccctggg         | 2698 |
| D | b | 2556 | GGCTCCTTCAACGACGCGGTGTACCGCGGGAACTCAAAACACTTGGGGCTCTTTGGGA           | 2615 |
| O | y | 2699 | ccalatalaagaagcagaagtltgaagalaatatacattgtaacttccagaatcgaagcctct      | 2758 |
| D | b | 2616 | CCCTACATTCAGAGCGGAAGTTGAAGAACATCATGTAATCTTTAAAAACCGAGCGCT            | 2675 |



QY 2759 cgtccctattctctctatctctagccttattcttaaggaagatcagaagcagaagca 2818  
|||||  
Db 2676 CGTCCCTATTCTCTTACTACTGAGCCTTATTCTTATCCGGATGATCAGAGCAAGGGGCA 2735  
QY 2819 gaacctgaagaaaaacttctgaagccttaagaaacaaacttacttcttggaagtcaa 2878  
|||||  
Db 2736 GAACCTGACCAACTTCGTCCACCAATAGAAACCAAGAACTTCTTGGAAATGACG 2795  
QY 2879 catcatatgagcccaactaaagatgagtttgactgcaagcctgggcttatctctgat 2938  
|||||  
Db 2796 CATCATATGGACCCACAGAAAGAGAGTGTGACTGCAAGCCTGGGCTACTTTTGTGAT 2855  
QY 2939 gttagacctgaaaaaagatgtgctactcaggccttgatctgaccctctgtctgcacact 2998  
|||||  
Db 2856 GTTACCTGGAAAAAGATGTGCACACTGAGCTTGATGGGCCCTCTGTGATCGCGCGCC 2915  
QY 2999 aacacactgaacctgtctcatatgagacaagtgtacagatcagaagaaattgctctgttttc 3058  
|||||  
Db 2916 AACACCTGAAAGCTGTCCACGGTAGACAACTGACCGTGCAGAAATTTGCTGTGTTTTC 2975  
QY 3059 accactcttgatgagacccaaagctggtgacttcaactgaaataatgaaagaacactgcag 3118  
|||||  
Db 2976 ACTATTTTGTATGAGACAAAGACTGTGACTTCACTGAAATGTGAAAGAACTGCGCG 3035  
QY 3119 gctccctgcaatatccagatggaagatcccactttaagaagaattatcgtcttcacatgca 3178  
|||||  
Db 3036 GCCCCTGCGACCTGAGATGGAGAGACCCCACTGAAAGAAACTATTCGTTCCATGCA 3095  
QY 3179 atcaatgtctacataatggaatacactactgctgtaagttaagtctcagaatcagaagatt 3238  
|||||  
Db 3096 ATCAATGGCTATGTATGATGATACACTCCCTGGCTTGTATGATGCTCAATCTAAAGGATC 3155  
QY 3239 cgaatgcatctgctcagaatggaagcaatgaaacatccatcttcatctatctcagttga 3298  
|||||  
Db 3156 CGATGGATATCTGCTCAGCATGGGCAAGATGAATAATATCCATTCCATTCTTTAGCGGA 3215  
QY 3299 catgtgtctactgtacgaaaaaagaagaagatataaatgtgacatgtacatctctatcca 3358  
|||||  
Db 3216 CACGTGTTCAGTGTACGGAAGAAAGAGAGATATAAATGCGCTGTACAACTCTATCCG 3275  
QY 3359 ggtgtttttgagaagaagtggaattttacatccaaagcttggaatttggcgggtgaatgc 3418  
|||||  
Db 3276 GGTGTCTTTTGAAGAGTGAAGATCTACCGTCCAAAGTTGGAATTTGGCAATTAAGAATGC 3335  
QY 3419 cttaattgagcagcatctacatgctcggagatgagcacaacttctctgtgtacagcaatag 3478  
|||||  
Db 3336 CTGATTTGGCAGACACTGCAAGCTGGGATGAGCACGACTTTCTTGCTGTACAGCAAGAG 3395  
QY 3479 tgtcagaactcccttgggaatgtctcttgagacatltagagaatttcagattacagcttca 3538  
|||||  
Db 3396 TGTGAGCTCCACTGGGAATGGCTTTCTGAGCAGCATTAAGATTTTCAGATCAGACTTCA 3455  
QY 3539 ggaacaaatgagcagtggtgcccacaaagctgagcagacttcatatccggatcaatcaat 3598  
|||||  
Db 3456 GGAGAGATATGAGACGTGGGCCCCCAAGCTGGCCAGACTTATTTATTCGGGATCAATCAT 3515  
QY 3599 gctctgagacccaagagagcccttctcttgagatcgaagtgtgctctgttgagccaaatgatt 3658  
|||||  
Db 3516 GCCTGGAGCAACCAAGGATCCCCACTCTCTGATCAAGGTGGATCTGTGTGGCAACATGATC 3575  
QY 3659 attcaaggcatcaagagcccaaggtgtcccgltcagaagttctccagccttcaatctctcag 3718  
|||||  
Db 3576 ATTCACGCGATCATGACCCAGAGGTGCCGTCAGAAAGTTTTCAGAGCTCTACATCTCCAG 3635  
QY 3719 ttatcatcatgtatagttctgtatggaagaagtgtgagacttcatcgagaatccact 3778  
|||||  
Db 3636 TTTATCTATCTATGTACATCTGTGAGGAGAACTGGCAAGATTTACCGGAGGATTTCCACG 3695  
QY 3779 ggaaccttaatgtctctcttggcgaatgtgataatcaltcgatlaanaacacaaatatttt 3838  
|||||  
Db 3696 GGCACCTTAATGTCTTCTTTGGCAATGTGAGCGCATGTGGGATTAACACAAATATTTT 3755

QY 3839 aacctccaattatgtctgatbaatccgttctgaccaccaactcatatagcatctgagc 3898  
|||||  
Db 3756 AACCTCCGATGTGGCTCGGATCATCCGTTTGGACCCAAACACATTAACAGATCCGAGC 3815  
QY 3899 actcttcagatggagttgatggctgtgatattaagtgttcacatgcatgcatggagatg 3958  
|||||  
Db 3816 ACTCTTCATGAGATGTATGGCTGTGATTTAAACAGTTTCACACATGCCCTGGGAATG 3875  
QY 3959 gagaataagcaatatcagatgcagaatcttactcttcatcttacttacaatatttt 4018  
|||||  
Db 3876 CAGATTAAGGATATTCAGACTCAGATTCAGGCGCTCTCTCCACACATTAATATTTT 3935  
QY 4019 gccacctgtctctcttcaaaagctcgacttcaactcagaaggaggaatgacctgaga 4078  
|||||  
Db 3936 GCCACCTGGTCTCTTTCACAAAGCCGACTTCACTCCAGGGGCGAGAAATGCTTGCGA 3995  
QY 4079 cctcagtgataataccaaaagatgtgctgtaagtgtgacttccagaagaacatgaaatgc 4138  
|||||  
Db 3996 CCCCAGGAGACAGCGAGAGAGTGGCTGAGGTGACGTGCAGAAAGCGTGAAGTTC 4055  
QY 4139 acaagagttaactcaaggagtaaatctctgtcttaccagaatgtatgtgaaggatgc 4198  
|||||  
Db 4056 ACAGGATCACACCCAGGCGGTGAAGTCCCTGCTACAGACATGTGTGAAGAGACTTC 4115  
QY 4199 ctcatctcagaagtcgaagatgagcaatcagtgaaactcttcttccagaatggaagta 4258  
|||||  
Db 4116 CTGCTGTCAAGTATGACAGAGCGCCCGCTGAGCCCTGTCTTTCAGAGCGGCACACG 4175  
QY 4259 aaggttttcagggaatacgaactccttcaacactgtggaactctctagaccaccg 4318  
|||||  
Db 4176 AAGGTTTTTCAGGGCAATCAGGACTCTTCACCCCGGGTGAAGCTGTGGACCCCG 4235  
QY 4319 ttactgactcgtactccttgcgaatcgaatcccccagaagtgtggtgacacagatgtccctgag 4378  
|||||  
Db 4236 CTGTTACAGCGCTACTGATGATCCACCCACAGAGCTGGGGGAGCAGCATGCCCTGAGC 4295  
QY 4379 atgagagttctgtgctgcgagcagaagcactctactgca 4417  
|||||  
Db 4296 CTCGAGGTTCTAGGATGTGAGGACAGAGATCTTACTGA 4334

RESULT 10  
LOCUS AR110041 4451 bp DNA PAT 14-FEB-2001  
DEFINITION Sequence 42 from patent US 6114148.  
ACCESSION AR110041  
VERSION AR110041.1 GI:12826317  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4451)  
AUTHORS Seed,B. and Haas,J.  
TITLE High level expression of proteins  
JOURNAL Patent: US 6114148-A 42 05-SEP-2000.  
FEATURES  
source 1..4451  
BASE COUNT 972 a 1568 c 1234 g 677 t  
ORIGIN

Query Match 58.9%; Score 2603.8; DB 9; Length 4451;  
Best Local Similarity 74.8%; Pred. No. 0;  
Matches 3262; Conservative 0; Mismatches 1097; Indels 0; Gaps 0;

QY 59 gccaccagaatactactcgtgtgagtgagactgtcatatggaactatgcaaatgtat 118  
|||||  
Db 84 GCCACCCCGCTACTACTGAGCGCGGTGAGACTGTCTGGGACTCATGACAGACGAG 143  
QY 119 ctggtgagcgtcgtgtgagcagaagattcctcctctagatgccaataatcttctcattc 178  
|||||  
Db 144 CTGGGCGAGTCCCGGTGAGACGCCCGCTCCCGCGTGGCCAGAGACTTCCCTTC 203



|    |      |   |      |
|----|------|---|------|
| OY | 179  | aacccctcaagtcgtgtatcaaaaaagactcgttttgtgaattcaagtggttaaccttlttaac    | 238  |
| Db | 204  | AACCAACGACGGCTGGTGTACAAAGAAAACCCGTTCGTGTGATGTACCGACGACCACTGTTTCAAC  | 263  |
| OY | 239  | atgcctaaagcaagaagcccaaccttgatgggttcgtcctgaagctccatcatcaagcttgaagtt  | 298  |
| Db | 264  | ATTGCCAAGACCCGGCCCCCTTGGATGGCTTCCTTG66CCCAACCATTCAGGCGCAAGGTG       | 323  |
| OY | 299  | taigatacagtygtlcaattacttaacttaagaacatggtcttccactctgtcaagttcaatgct   | 358  |
| Db | 324  | TACGACACCGTGGTATCACCTCTGAAGAACATGCGACGCCACCCGTAGCTGCACGCC           | 383  |
| OY | 359  | gttggtgatccctactcaggaaagcttctgggagctcgaatatgatalgatalcagaaccaagttaa | 418  |
| Db | 384  | GTGGGGCTGAGCTACTCTGGAAGGCGACGAGGGGGCCGATGACGACGACGCTCCAG            | 443  |
| OY | 419  | agggagaagaagaatgataaaagctctcccttgtytgaagcacaacatalgtctcgaagtc       | 478  |
| Db | 444  | CGCAGAGAAGGAGGACGACAGGTGTTCGGGGGGGAGCCACACCTAGGTGGGAGGTG            | 503  |
| OY | 479  | ctgaagaagaatgltccaatgagctctgaaccaactgtgccttaactactatalacttct        | 538  |
| Db | 504  | CTTAGAGAGAACGCGCCTTAGGCGACGACCCCTGTGCTCACTCAACGACTACTGAGC           | 563  |
| OY | 539  | catygacccttggttaaaagactcgtgaattcaagagcctatttggagccctactagtatga      | 598  |
| Db | 564  | CACGTGGACCTG6TGTGAAGATCTGAACACGCGGCTATGTGGCGCTGTGTGTGTGC            | 623  |
| OY | 599  | gaaggagatcttgcgccaaggaagaagacagaccttgacaaattatatactacttttgt         | 658  |
| Db | 624  | GAGGGCGAGCTGGCCCAAGGAGAAACCCAGACCCTTCACAACTTCATCTGTGTTGCC           | 683  |
| OY | 659  | gtatttgatgaagggaagaagtgtgcactcaagaacaagaactccttgatgcagatag          | 718  |
| Db | 684  | GTGTTCCACAGAGGGGAGAGACTGTGCACGACGAGACTTAAACAAAGCCTATGACGAGCAGC      | 743  |
| OY | 719  | gatgtcattctctcgggagcttgacctaaatgacacagtcataatggtatgtatlaaacag       | 778  |
| Db | 744  | GAGCGCGCCACGCGCGCGCTGTGGCCCAAGATGGTCAACCGTTAAGGGGTACGTGAACGCC       | 803  |
| OY | 779  | tctctgcgaagtcctgattgtgatgcacaggaagaactcaagtcattatitgcatgttggaagt    | 838  |
| Db | 804  | AGCCTTGCCGCGCTGTATCGGCTCCACCGCAAGGCGTGTACTGTGACGTACTATCGGCATG       | 863  |
| OY | 839  | ggcaaccaactcctaaagtgcactcaatccttcctcgaagtcacacaattccttbtgtggagac    | 898  |
| Db | 864  | GGCAGCACCCCTAGGGTGCACACATCTTCTGTGAAGGCCAACACTTCTGTGTGGGCAC          | 923  |
| OY | 899  | catgcgcagcgctccttggaaatctgcgcaataactttccttaactgtgtcaaacacttgg       | 958  |
| Db | 924  | CACGCGCAGGCGCAGCCTGGAGATCACCCCAATCACCTTCTCACTAGCCAGACCCCTGCTG       | 983  |
| OY | 959  | atggaccttggaagtttactactggttttgcatactctcccaacaacatgagtgaagt          | 1018 |
| Db | 984  | ATTGACCTAAGGCCAGTTCTCGTCTGTTCTGCCACATTCAGCAGCCACCAAGACAGCGCATG      | 1043 |
| OY | 1019 | gaagcttatgtlcaaaatlagacagctgtlccaagaggaaccccaactcagaatgaanaatat     | 1078 |
| Db | 1044 | GAGGCTTAAGTGAAGGTGGACAGCTGCCCGGAGAGACCCCACTGCGCATGAAGAACAC          | 1103 |
| OY | 1079 | gaagaagcggaaagactatgatgtatgtacttaactgtatcttgaatgtagtgtgtaagttt      | 1138 |
| Db | 1104 | GAGGAGGCGCAGGACTTACGACGACGACCTGACCCGAGAGAGGAATGATGTGTTAGCTTC        | 1163 |
| OY | 1139 | gatataataaacctctcttccttaataatgcctagtgagtttgcgaagaagaatcctaa         | 1198 |
| Db | 1164 | GAGGAGGACAAACAGCCCACTTCAATCCAGATTCGAGGCTGGCCAAAGACACCTTAAG          | 1223 |
| OY | 1199 | acttggttacattacattgctctgtcgaagaaggaagactggagactatgtctccctagttctc    | 1258 |
| Db | 1224 | ACCTGGGTGTGACTATACATCGCCCGGAGGAGAGATCTGGGATGACGCCCGCTAGTATTGT       | 1283 |

|    |      |   |      |
|----|------|---|------|
| QY | 1259 | gccccgatacgcaagattataaagtcaatatttgaacaatgcccccaagcatttgt          | 1318 |
| Db | 1284 | gcccccgacgacccaccacttccaaagccattacttaccmaacggcccccagcgctcggc      | 1343 |
| QY | 1319 | aggaaagtcaaaaagatccgattatbgaatacaagatggaacctttaagatcgtttaa        | 1378 |
| Db | 1344 | gcgaagtacaaagagtgccgctttatbgtctaacccgacgaacattttcaamgacccgcag     | 1403 |
| QY | 1379 | gcatatcagaatgaataaggaatctbvgaaaccttacttlatggygaagtbgaaacaa        | 1438 |
| Db | 1404 | gccattccagcagcagatccggcatctctgagccctctcttaccggcgagstggcgacacac    | 1463 |
| QY | 1439 | ctgtgtattatalttaagaaatcaagaagcagaacatatcaatccctcaegnaatc          | 1498 |
| Db | 1464 | ctgctgattcatctttcaagaacccagggcagagcccttaacacattaccctccacggcatt    | 1523 |
| QY | 1499 | actatgccccctctgtgttttcaagaagttaccbaaagtgyltaaacattttgaatc         | 1558 |
| Db | 1524 | accgacggccccccctgtatagccggccttcccaabggggtgaagacctgaagac           | 1583 |
| QY | 1559 | ttccaattctgcagaggagaataatcaaatataaattgacagtgactbgtlaaagaatgg      | 1618 |
| Db | 1584 | ttccccattcctggcccgcgagattcttcaatataaattgacacgctgacccgtgagagacgc   | 1643 |
| QY | 1619 | ccaactaaatcagaatctctcggtgcctgaccggctattacttactgttgttaatatggag     | 1678 |
| Db | 1644 | cccaccagaagcagacccccctcgctctacccgcttaccagcagcttctgtmaacattgag     | 1703 |
| QY | 1679 | agagatctagcttcaagatctccattgtgcctctctcaatctgtcaaaagaatctgtagat     | 1738 |
| Db | 1704 | cgccacctggccttcggagacgattctggccccctctctatctgttaaaagagacgttgac     | 1763 |
| QY | 1739 | caaaagagaacacgaataatggtcaagaagaagaaigtatctctgttctgtattgat         | 1798 |
| Db | 1764 | cagcgcgacacacgattcatatgacgacacacgacaaagtattctgttcaacgctgtttcac    | 1823 |
| QY | 1799 | gagaaacgaagctgtgtaacctacagaagaatatacaagcttctctcccaatccagctbga     | 1858 |
| Db | 1824 | gagaaacgcgactggtatgttgcacggaacatcatcagcgttctctgcccmaaccccgctggc   | 1883 |
| QY | 1859 | gtgcagcttgaagatccaaaggttccaaagctccaacatabgacaaagcaatgaatgctat     | 1918 |
| Db | 1884 | gtgcagcctggaagatcccgaggttccagggccaggaacatattcacagacatcmaacggctac  | 1943 |
| QY | 1919 | gtttttgtatgttgaagttgtcaagttgttgacgagtgagacatcgtgtacatctla         | 1978 |
| Db | 1944 | gtgtttcagacaccttgacagctgacgctgtgctgtcagatgagagtgctactgctgacatcctg | 2003 |
| QY | 1979 | agcatctgagacagactgaactctctcttctgtcttctctctgtgalatalaacttcaaacac   | 2038 |
| Db | 2004 | agcatctggcgcccgacgcgacttctctgaggggtttcttccggggattaccttcaagcac     | 2063 |
| QY | 2039 | aaaatggtctatgtgaagacacactcaacctatccacttctcaagsgaagactgtctctatg    | 2098 |
| Db | 2064 | aaagtgtgtatcgaggaacacctgacccgtttcccttccgggagagactgtgtttatcg       | 2123 |
| QY | 2099 | tccagttgaaaaacccaagttatgtatcttcggggtgacaaactcaagacttccgttgaaacga  | 2158 |
| Db | 2124 | ttctatggagaaacccccgctgttgattttctggctgcccacaaacagcgacttccgcacacgc  | 2183 |
| QY | 2159 | ggaatgaacgccttaccgaaggttctctgtgtgacaagaagaacactggtatattacgag      | 2218 |
| Db | 2184 | ggcatgactgccctgctgaagatctccagctggcgaacaaacacggcgactactgacgacg     | 2243 |
| QY | 2219 | gaacgttatgaagatatttcagcatacttgcgtgagtaaaacaatgccaattgaaccaaga     | 2278 |
| Db | 2244 | gacagctacgaggaacatctccgctactcctgctgttccaaamaacagccatgacggccgc     | 2303 |
| QY | 2279 | agcttctccagaatbtaagaacacttaagacataggcaaaacaaatbtaatgtgcaaccca     | 2338 |
| Db | 2304 | ttcttttctccaaaactcccgccaccccaagcgcgtcgaabaacgaatttaacgccaccccc    | 2363 |
| QY | 2339 | ccagttcttgaaacccaatacaagggaataaactcgtactactcttccagttcaagataaag    | 2398 |

Db 2364 CCGGTGTAAGCCGACACGAGATGATCAACCCGACACCTGTGCAAAAGCAGCAGAG 2423  
OY 2399 gaaattgacatgtatgatacatatcagatgtgaaatgaaaggaatttggatactat 2458  
Db 2424 GAGATGACATGACGACGACACCATCAGCTGAGATGAAGAAGAGAGACTTGCACATTCAC 2483  
OY 2459 gatgagaaatgaatcagagcccgcaagcttcaaaaagaacaagaactattat 2518  
Db 2484 GAGGAGGAGCAACACGAGACCCCGCTCTTCCAAAAGAAACCCGCACTACTTCATC 2543  
OY 2519 gctgagatgagagagcttggatgtatgagatgagatcccaatgltctaaagaac 2578  
Db 2544 GCGGCGTGAAGCCCTGTGGACTACGCGATGAGCACACCCCGACGCTCTGGGACAC 2603  
OY 2579 agggctagaagtgagatgtccctcagatcgaagaagtgcttctcagaatcttcat 2638  
Db 2604 CCGGCCAGAGCGGACGCTCCCGAGTTCAAGAAAGTGTTGCCAGAGTTCCACGAC 2663  
OY 2639 gactccttactcagccctatataccgtggagaactaaatgaacatttggactccg 2698  
Db 2664 GCGAGCTTCAACCCGACCCCTGTACCGCGGCGAGCTGAAGAGACACTGGGCTCTCGC 2723  
OY 2699 ccataataagagagagatgtgaagataatatactgtgaacttcaagaatcagccct 2758  
Db 2724 CCGTACTCCGCGGAGGTGAGACATCATGTGATGCTTCCGACACCAAGCCTTC 2783  
OY 2759 cgtccctacttctctatctcagcttacttctatgagagatcagaggaagagaca 2818  
Db 2784 CGGCGCTACTCTCTTACTCTCTCTGTATGATGATGATGATGATGATGATGATGATG 2843  
OY 2819 gaacctgaagaaacttctgaagcctaaatgaacaaacttacttggaaatgaca 2878  
Db 2844 GACCCCGGCAAGACTTGTGAAGCCCAACGACCTTAAGACTTCTGGAAGTGCGAG 2903  
OY 2879 cactatataagcaccataagatgattgactgcaagaagcttggacttacttctat 2938  
Db 2904 CACCAATGAGCCCGCCACGAGAGAGATGATGATGATGATGATGATGATGATGATG 2963  
OY 2939 gctcagctggaagaaatgtgcaatcagcctgatttggaccctctgctgctcacaat 2998  
Db 2964 GTGAGCTGAG 3023  
OY 2999 aacacactgaacctgtcatatggagagaagtgacagtaagaatgtgctgttctc 3058  
Db 3024 AACACCTGTAACCCCGCCACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3083  
OY 3059 accatcttgatagagacaaagctgtgacttactgtaaaaatgtaaaaagaactgag 3118  
Db 3084 ACCATCTTCAACGAGACTTAAGAGCTGTGATTCACCGAAGACATGAGCGCACTGCCG 3143  
OY 3119 gctccctgcaatatacagatggaagatcccaactttaaagaatatacgtctccatga 3178  
Db 3144 GCCCGCTGCAATCCAGATGGAAGATCCACCTTCAAGAGAACTACCGCTCCACGCC 3203  
OY 3179 atcaatggtcaataatgataactaactggttagtaatgtgctcagatcacaagaatt 3238  
Db 3204 ATCAACGGCTACATCATGAGACACCTGCCCGCTGTGATGAGCCAGACACAGGCAATC 3263  
OY 3239 cagatgatactgtcagcagatggagcagaatgaacatcacttctatcattcagtgga 3298  
Db 3264 CCGTGTGACTGCTGTATGTGGCAGCAACGAGACATCCACAGCATCTTCAAGCGCG 3323  
OY 3299 catgtgtcactgtacgaaagaaagagatataaagtgaactgtacatctcatatca 3358  
Db 3324 CAGGTTTCAACGCGGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3383  
OY 3359 ggtgttlttgaagacatggaatgttaccatccaaagcttgaatttggcgggtggaatgc 3418  
Db 3384 GCGGTTTCAACGAGACTGTGAGATGCTGCGCAGCAAGCGCGGAGACTGCGCGTGAAGTGC 3443  
OY 3419 ctatattggagacatcactatgctgagatgagcacttcttctgtgttaagaacataatg 3478  
Db 3444 CTGATGGCGAGACCTGACCGCGGATGAGACACCTGTTCTGTGTATGAGCAACAG 3503  
OY 3479 tgcagactcccttggagatgcttctgacacatlagatatttccagatlaacgttca 3538  
Db 3504 TGCCAGACCCCGCTGGGATGGCCAGGGGCGCACATTCGCGAGCTTCCAGATCCGCGAGC 3563  
OY 3539 ggaacataatgagcagtggtcccaaaagcttggcagacttcaatctccagatcaat 3598  
Db 3564 GCGGAGTACGGCGAGTGGGCTCCCAAGCTGGCCCGCTGCACTACAGCGCGAGCATCAG 3623  
OY 3599 gcttggagcagcagagaccccttcttggatgaatgtgagatgtgtgtggaacatgatt 3658  
Db 3624 GCTGTGTGACCAAGAGACCTTCTCTGTGATCAAGGTGACCTGTGGGCCCCCATATC 3683  
OY 3659 attcagagcatcaagaccagagtgccgctagaagatcttcccaagcttcaatctcag 3718  
Db 3684 ATCCAGCGCATCAAGACCGAGGGGCGCGCAGAAAGTTCAACACCTGTACATCAGCCAG 3743  
OY 3719 ttatcatatgtatgacttctgagatggaagaagtggcagacttccaggaatctcact 3778  
Db 3744 TTCAATCATGATGATCTCTAGAGGCGCAAGAAAGTGGCAGACCTACCGCGCAACAGCAC 3803  
OY 3779 ggaacctaatgttcttcttggcaatgtgattcatctggaataaacaacatattt 3838  
Db 3804 GGCACCTGATGTGTTTCTTGGCACTGAGACAGACGCGCATCAACACATCTTC 3863  
OY 3839 aacctccaattatgtctgaatacatccgttggaccacaactaliaatgactcagc 3898  
Db 3864 AACCCCGCATATGCGCGCTGATCCGCTGACCCGCTGACCCGCTGACATCAGATCCGAGC 3923  
OY 3899 actcttgcagatgagatgtgagctgtgatttaattgatttgcagatgcaattggagatg 3958  
Db 3924 ACCCTGCGCATGAGAGATGATGGCTGCACTGAAGACCTGACACATGACCTCGGAGATG 3983  
OY 3959 gagatgaagcaataatgaatgacagattactgtcttcaacttcaacttcaatattt 4018  
Db 3984 GAGAGCAAGGCCATCAG 4043  
OY 4019 gcaacctgttctctcaaaagctcagacttccactccaaggaagatgaatgtctgaga 4078  
Db 4044 GCCACCTGAGACCCCGAG 4103  
OY 4079 cctcagtgtaataatccaagaagtgctgcaatgagatgtcagaagaacatggaatgc 4138  
Db 4104 CCCCAGGTAACACCCCAAG 4163  
OY 4139 acaggaatgaactcagagagatgaatctctgttctcagagatgtaatgtgaagagat 4198  
Db 4164 ACTGGCGTACACCGAG 4223  
OY 4199 ctcatctcagcagatcaagatgagcactcagatgtgactcttcttcaagaatggaagta 4258  
Db 4224 CTGATACAGACGACGAG 4283  
OY 4259 aaggttctcaggaataatcaagactcttcaacacttgttgaactctcagaaccacg 4318  
Db 4284 AAGGTTTCAAGGCGCAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4343  
OY 4319 ttactgaactgctacttcaatccaccccaagtggtgtgagcagagatgtcagag 4378  
Db 4344 CTGCTGAGCCGCTACTCTGCGATCCACCCCAAGAGCTGGGTGACACGATGCGCTCGC 4403  
OY 4379 atgagatctggtgctgagagcagagacacttactga 4417  
Db 4404 ATGAGAGTGTGGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4442

RESULT 11  
AR003710 AR003710 9009 bp DNA PAT 04-DEC-1998  
DEFINITION Sequence 1 from patent US 5744446.  
ACCESSION AR003710  
VERSION AR003710.1 GI:3964969

|            |        |                                     |
|------------|--------|-------------------------------------|
| KEYWORDS   | .      | Unknown.                            |
| SOURCE     | .      | Unknown.                            |
| ORGANISM   | .      | Unclassified.                       |
| REFERENCE  | 1      | (bases 1 to 9009)                   |
| AUTHORS    |        | Lollar,J.S. and Runge,M.S.          |
| TITLE      |        | Hybrid human/animal factor VIII     |
| JOURNAL    |        | Patent: US 5744446-A 1 28-APR-1998; |
| FEATURES   |        | Location/Qualifiers                 |
| SOURCE     | 1.     | 9009                                |
| BASE COUNT |        | /organism="unknown"                 |
| ORIGIN     | 2853 a | 1907 c 1844 g 2405                  |

|                            |        |              |          |             |
|----------------------------|--------|--------------|----------|-------------|
| Query Match                | 52.9%  | Score 2336.4 | DB 9     | Length 9009 |
| Best Local Similarity      | 100.0% | Pred. NO. 0  |          |             |
| Matches 2337; Conservative | 0      | Mismatches 1 | Indels 0 | Gaps 0      |

|    |     |  |     |
|----|-----|--|-----|
| OY | 1   | catgcaataaagctctccacccgctctcttcgtgctcttcgagttcgtcttagtc            | 60  |
| Db | 150 | CATGCAAAATAGAGCTCTCCACCTGCTCTTCTTGCTGCTTTTCGATTCGCTTTAGTGC         | 209 |
| OY | 61  | caccagaagataactacctgggtgagctggaactgtcaatggactatatagaagatcat        | 120 |
| Db | 210 | CACCAGAAGATACTACTCTGGGTGAGTGGAACTGTCTATGGGACTATATATGCAAGTATCT      | 269 |
| OY | 121 | cagtgagctgcctctgtgacgcgaagattctccctcctagatgtgcacaaatcttccattcaa    | 180 |
| Db | 270 | CGGTAGCTGCTGTGGAGCCAGAGATTTCCTCTAGAGTSCCAAAATCTTTCCATTCAAA         | 329 |
| OY | 181 | caactcagtcgtgtacaaaaaagactcgttttgtagaatcaaggttcacacttcaacat        | 240 |
| Db | 330 | CACCTCAGTGGTGTACAAAAAGACTCTGTTTGTAGAAATTCACGGTTCCACTTTTCAACAT      | 389 |
| OY | 241 | cgctcaagccaaagccaccaccctggatgggtctgcgtcagatgccacacatccagctgagttta  | 300 |
| Db | 390 | CGCTAAGCCAAAGGCCACCCCTGGATGGGTCTGTAGTCTTACCATCAGGCTGAGGTTTA        | 449 |
| OY | 301 | tgatacagtggtcatltaacttaagaacaatgagcttcccatccgttcagttcattcatgct     | 360 |
| Db | 450 | TGATACAGTGGTCTTACCTTTAAGAACATAGGCTCCCATCCGTGACGTCTTATGCTGT         | 509 |
| OY | 361 | tgtgtatcctactctggaagctctctctgagggagctgtgaatgaigtatcagaaccaatgaag   | 420 |
| Db | 510 | TGTGTATCCTACTACGGAAACCTTCTGAGGAGCTGAATATGATGATCAGACCCAGTCAAG       | 569 |
| OY | 421 | ggagaagaagatgataaagtcttccctctggtggaagccataactatgctctgagagtcct      | 480 |
| Db | 570 | GGAGAAAGAAAGATGATTAAGCTTCTTCCCTGGTGGAGCCATATATGCTCGGAGGTCCT        | 629 |
| OY | 481 | gaagagaaatggtccaaatggcctcttgaccacacgttccttaactactaatcttctca        | 540 |
| Db | 630 | GAAAGAAATGTGTCANATGGCTCTGTGACCACACGTGTGCTTACCTACTCATATCTTTCTCA     | 689 |
| OY | 541 | tgtgagacctgtgtaaaagacttgaattcaagcctcatcttgagccctactagatgtgaga      | 600 |
| Db | 690 | TGTGTGACCTGTAAAGACTTGAATTCAGGCTTCATTTGAGCCCTACTGATGTAGTGAAGA       | 749 |
| OY | 601 | agggagctctggccaaggaaaaagacaacagacactgtgacaaattataactacttctgtgt     | 660 |
| Db | 750 | AGGGAGCTCTGGCCMAAGAAAAAGACACAGACTTTCACAAATTAATATCTACTTTTTCCTGT     | 809 |
| OY | 661 | atttgatgaaaggaagaagtgtgagcactcagaagaacaaagaactcttataatcagataggga   | 720 |
| Db | 810 | ATTTGATGAAGGGAAGAAAGTTGGCACTCAGAAAPAAAGAAATCTCTTATGACGAGATVGGGA    | 869 |
| OY | 721 | tgctgcatactgctcgggcctgagcctaataatgacacacagtcatacgtgtatgtataacaggtc | 780 |
| Db | 870 | TGCTGCATCTGCTGGGCCCTGGCCCTAAATATGCACACAGTCAATAGCTTATGTAAACAGTTC    | 929 |
| OY | 781 | tcctgcaggtctcattgagatgacacaggaatcactatctatctgcatgtgattggaaatggg    | 840 |

|    |      |   |      |
|----|------|---|------|
| Dd | 930  | TCGCGACGGTTGATTGGATGCCACAGAAATCAGTCCTATTGGCATGTGATTGGAATGGG       | 989  |
| Qy | 841  | caccactctctgaagtgcaactcaatattccctcgaaggtcacaactctctgttgaaggacaa   | 900  |
| Dd | 990  | CACACCTCTGAAGTGACACTCAATATTCTCGAAGGTCACACTTCTCTGTGAGGACCA         | 1049 |
| Qy | 901  | tcgccaggcgtcccttgganaalctgcacaatacttccctactctgctcaacaactctgat     | 960  |
| Dd | 1050 | TCGCGACGCGTCTCTTGGAAATCTCGCCAAATACCTTCCCTACTGCGCAACACTCTGTAT      | 1109 |
| Qy | 961  | ggaccttggacaagtttctactggtttgtcaatctcttcocacccaacatgatggatga       | 1022 |
| Dd | 1110 | GGACCTTGGACAAGTTTCTACTGTTTGTCTATTCCTTCCACCAACATGATGGCATGGA        | 1166 |
| Qy | 1021 | agcttatbctcaaattagacagctbctccagaaggagcccaactacgaatgaanaataatga    | 1086 |
| Dd | 1170 | AGCTTATCTCAAATAGACAGCTGTCCAGAGGAAACCCCACTACGATATAAAATATATGA       | 1222 |
| Qy | 1081 | agaagcggnaagactatgatgatctctactgaattctgaatltgaatltgtgtcagglttga    | 1140 |
| Dd | 1230 | AGAAGCGGAACACTATGATGATGATCTTACTGATTCCTGAAATGATGATGTGACGGTTGA      | 1289 |
| Qy | 1141 | tgatgaacaactctctctcccttaltccaatctgcctcaagtgtgccagaagatcctcaaac    | 1200 |
| Dd | 1290 | TGATGACAAACTCTCTCTTCTTATCCAAATTCGCTACGATGGCCAGAAACATCCATAAAC      | 1349 |
| Qy | 1201 | ttgggttaacttaactbctctctctaagaaggagactbaggatagctcccttagtcctgcg     | 1266 |
| Dd | 1350 | TTGGGTACATTACATTCCTCTCTGAAAGGAGGACTGGACATAGCTCCCTTAGTCTCGC        | 1409 |
| Qy | 1261 | ccccgatcacgaagattataaagltcaaltlttgaacaaltgcccccaacgaagattggtag    | 1320 |
| Dd | 1410 | CCCCGATGACAGAGATTATAAAGTGCAATTTTGAACAAATGGCCGCACAGGATTTGTAG       | 1466 |
| Qy | 1321 | gaagtacaaaaagtcgccgattatgtgcatacacagatlgaaaccttaagaactcgtgaagc    | 1386 |
| Dd | 1470 | GAAGTACAAAAAAGTCCGATTATTAGGATACACAGATGAAACCTTTAAACACTCGTAGAAC     | 1529 |
| Qy | 1381 | tattcagcatgaatcacgaactctgtggagaccttaacttttggggaaagttggagacaact    | 1440 |
| Dd | 1530 | TATTTCAGCATGAATCAGGAATCTTTGGACCTTTACTTTATGGGGAAGTTGGAGACACT       | 1589 |
| Qy | 1441 | gttattatataattagaatcaagaacagacagacataataacatccctcacaagatcac       | 1500 |
| Dd | 1590 | GTTGATTATATTTAAGAATCAGACAGACACCATATTAATCTACCTCCACAGGAATCAC        | 1649 |
| Qy | 1501 | tgatgtccgtcccttgttattccaaggagattaccaaaagltgaanaacatttgaagaattc    | 1560 |
| Dd | 1650 | TGATGTCCGTCCTTTGTATTCAAGGAGATTACCAAAAGGTGTAAAAACATTTGAAAGATT      | 1709 |
| Qy | 1561 | tcgaattctgcaggaganaaltattcaanaataaagtgaagtgagctgttgaagaatgggccc   | 1620 |
| Dd | 1710 | TCGATTTCTGCGAGAAATATTTCAAATTAATTAATGACAGTGACTGTGAAGATGGGCC        | 1769 |
| Qy | 1621 | aactcaatcagatcctcgtggtcccttgagcccgctatctctctagtcttcgttlaatatgagag | 1680 |
| Dd | 1770 | AACATAATMCASATCTCTGGTGCTGACCCGCTATTTACTGTACTTTGCTTAATATGAGAG      | 1829 |
| Qy | 1681 | agatctagcttcagactcatgtggccctcctccatctctgtctacaagaatcctgtatgatca   | 1740 |
| Dd | 1830 | AGATCTTAGCTTCAGACTCATTTGGCCCTCTCCATCTCTCAAAAGATCTGTGTAGATCA       | 1889 |
| Qy | 1741 | aaggagaaacagataatgtcacaagaagggaatgtcatccctggtttctcgtattgtatga     | 1800 |
| Dd | 1890 | AAGAGGAACCAAGTATATGTCTACACAAGAGGAATGTCACTCTGTTTCTGATTTTGAAGA      | 1949 |
| Qy | 1801 | gaacccgaagctgtgtaccttccacagagaatatlacacgcttctccccaactcagctgtgagt  | 1860 |
| Dd | 1950 | GAAACCGAAGCTGTACTCTCACAAGATATACAAAGCCTTTCTCCCAATCCAGCTGGAGT       | 2009 |
| Qy | 1861 | gcagctttagagatccagagtttccaagcctccaacatcatgtcaacagatcatalgtgtactgt | 1920 |
| Dd | 2010 | GCACCTTGAGGATTCAGAGATTTCCAAACCTCCCAACATCATGACACAGATCAATATGGCTATGT | 2069 |

|    |      |  |      |
|----|------|--|------|
| QY | 1921 | ttcttgtaagtgttcgaattgtctcaatttcttgatgagtggtgcatactcgtgatacttctaa | 1980 |
| Db | 2070 | TTTTTGATAGTTTCCAGTTCAGTTGTTTGGCAAGAGGTGCATACGTGTAATTTCAAG        | 2129 |
| QY | 1981 | catggaagcaacagacgaactcctctctcgtctctctcgtgataacctcaacaaca         | 2040 |
| Db | 2130 | CATTGAGCACACACTGACTTCTTTTGTCCTTCTTGATATACCTTCAAAACACA            | 2189 |
| QY | 2041 | aatgtctatbgaagacacactccacccatcccatctcgaagagaactgtctcatgctc       | 2100 |
| Db | 2190 | AATGTCTATGAGAGACACTCAACCTATTCCCATCTCAAGAGAAACTGCTTCATGTC         | 2249 |
| QY | 2101 | gattggaaaacccagctctatggaattctgggggtgcacaactcagaacttcggaaacaga    | 2160 |
| Db | 2250 | GATGGAAAAACCCAGGCTATAGGATTCTGGGGGTGCACAACCTCAACCTTCCGAACAGAGG    | 2309 |
| QY | 2161 | catgaacgcctactcgaagttctctagtttgacaagaacacgtgtgatattacacgga       | 2220 |
| Db | 2310 | CATGACCCGCTTACTGAGGTTTCTACTTGTGACAAACACTGCTGATTTATTGCAAGA        | 2369 |
| QY | 2221 | cegtatagaagatatctcagcatactctgctgagtaaaacaatlgccatlgacaacaaga     | 2280 |
| Db | 2370 | CAGTTATGAAGATATTTACACATACTGCTGAGTAAAAACAATGCCATTGAAACCAAGAG      | 2429 |
| QY | 2281 | ctctctcccgaaattcaagacacccctagcactaggcaaaagcaattatctgcacccca      | 2338 |
| Db | 2430 | CTTCTCCCAAAATTCAGACACCTTACCTATGGCAAAAGCAATTTATGCCACACCA          | 2487 |

| LOCUS                 | AR029065  | 9009 bp                   | DNA          | PAT    | 29-SEP-1999 |
|-----------------------|---|---------------------------|--------------|--------|-------------|
| DEFINITION            | Sequence  | 1 from patent US 5859204. |              |        |             |
| ACCESSION             | AR029065  |                           |              |        |             |
| VERSION               | AR029065.1  | GI:5941038                |              |        |             |
| KEYWORDS              |   |                           |              |        |             |
| SOURCE                | Unknown.  |                           |              |        |             |
| ORGANISM              | Unknown.  |                           |              |        |             |
| REFERENCE             | Unclassified.   |                           |              |        |             |
| AUTHORS               | 1 (bases 1 to 9009)   |                           |              |        |             |
| TITLE                 | Lollar/J. S.  |                           |              |        |             |
| JOURNAL               | Modified factor VIII  |                           |              |        |             |
| FEATURES              | Patent: US 5859204-A 1 12-JAN-1999;                                 |                           |              |        |             |
|                       | Location/Qualifiers   |                           |              |        |             |
|                       | 1..9009   |                           |              |        |             |
|                       | /organism="unknown"   |                           |              |        |             |
| BASE COUNT            | 2853 a  | 1907 c                    | 1844 g       | 2405 t |             |
| ORIGIN                |   |                           |              |        |             |
| Query Match           | 52.9%; Score 2336.4;  | DB 9;                     | Length 9009; |        |             |
| Best Local Similarity | 100.0%;   | Pred. No. 0;              |              |        |             |
| Matches 2337;         | Conservative 0;   | Mismatches 1;             | Indels 0;    | Gaps   |             |
| Qy 1                  | catgcaaatagagctctccaccctgctcttcctctgctgcctttagcattcgcattagtcg       | 60                        |              |        |             |
| Db 150                | CATGCAAAATAGAGCTCTCCACCCTGCTTCTTCTGTGCTCTTTTGGGATTCGCTTTACTGC       | 209                       |              |        |             |
| Qy 61                 | caccagaagatactacctggtgtgcagtggaactgcatggaactatagtcaagtgatct         | 120                       |              |        |             |
| Db 210                | CACCGAAGATACATCACTCGGGGTGAGTGAACCTGCTCAATGGACATATGCAAAAGTGATCT      | 269                       |              |        |             |
| Qy 121                | cgggtgaagctgctctgtagagcgaagaattccctcctcctagaagtgccaaatcctttccattcaa | 180                       |              |        |             |
| Db 270                | CGGTGACTGCTGCTGTGGACGCAAGATTTTCCTCTAGAGTGGCAAAATCTTTTCCATTCAA       | 329                       |              |        |             |
| Qy 181                | caacctcagctgtgtacaacaaagactctgtttgttagaattcaacggttcacctttcaacat     | 240                       |              |        |             |
| Db 330                | CACCTCACTCGGTGACAAAAGACTCTGTGTGTGAATTCACGGTTCACCTTTCAACAT           | 389                       |              |        |             |
| Qy 241                | cgtctagcgaagcgccaccttgatagggtctgtcttaggttcctaccatccagctgaggttta     | 300                       |              |        |             |

|    |      |  |      |
|----|------|--|------|
| Db | 390  | CGCTAAGCCAAAGGCCACCCCTGATAGGGCTGCTGAGTGGCTACCATCCAGGCTGAGAGTTA   | 445  |
| Qy | 301  | tgaatacagttgcaatlaacttaagaacatgagctcccaatccatgctacgtctcatgctgct  | 360  |
| Db | 450  | TGATACACTGGTCTATTACACTTAAGAACATAGGCTCCCATCTGTCAGTCACTCATGTGT     | 509  |
| Qy | 361  | tgtgtatcccttaactcggaaagctctcttgaggagcttgatatagatatacaaacagctcaag | 420  |
| Db | 510  | TGGTGTATCTACTGCTGGAAGGCTCTTGAGGGAGCTGAAATATGATGATCAACACAGTCAAG   | 569  |
| Qy | 421  | ggagaaaagaatataaagctctcccttggttgaaagcaataatgcttggaagctc          | 480  |
| Db | 570  | GGAGAAACAATATATTAAGTCTTCCCTGGTGGAAAGCATACATATGTGTGGCAGAGTCT      | 629  |
| Qy | 481  | gaaagaagatggtcccaatgagcctctgaacccactgtagcttaactactatactcttcca    | 540  |
| Db | 630  | GAAGAAGATGTCCTCAATAGGCTCTGACCCACATGTCGCTTACTACTCATATCTTTCA       | 689  |
| Qy | 541  | tgtgaccttggtlaaagaactgaaattcaagcctccatttgagccctactagtataaga      | 600  |
| Db | 690  | TGTGCACTTGTTAAAGCTTTGAATTCAGGCTCATTTGGAGCCCTATGATGTATGAGGA       | 749  |
| Qy | 601  | aaggagctctggccaaggaaaagaacagagaccttgcaaatataactctttgctgct        | 660  |
| Db | 750  | AGGAGGTGGCCCAAGAAAAAGACACAGACTTGCACAAATTATATCTTTTGTGCT           | 809  |
| Qy | 661  | atttgaatgaagggaagaatttggcaatcagaacaaaagaactccttgaatgcaagata      | 720  |
| Db | 810  | ATTGATGAAAGGAAAGATTGGCACTCAGAAACAAAGAACTCCTTGATGACAGATAGGGA      | 869  |
| Qy | 721  | tgtctgcatctgctctgggacctggccttaaaatgacacaaatgctatggttatataaagct   | 780  |
| Db | 870  | TGCTGCATCTCTCTGGGCTGGGCTTAATAATGACACAGTCAAGGTATATGAACAGTCT       | 929  |
| Qy | 781  | ctctgcaggtcttgaatgtagtcccaagaaataagctatcttgcaatgtaattggaatgg     | 840  |
| Db | 930  | TCTGCAGAGTCTGATTTGGATGGATGCCACAGAAATATAGTCTATTTGGCAGTATTTGGAATGG | 989  |
| Qy | 841  | caccacctctgaagtgcaactaaatctctcgtgaaggtlcaacaatctcttcttgyaagacca  | 900  |
| Db | 990  | CACCACTCTGAAATGCACTCAATATTTCCGGAAGGTCAACATTTCTGTGAGAACCA         | 1049 |
| Qy | 901  | tcgccaagcgtctcttggaatatctcgccaataactcttccttaactgctccaacactctgat  | 960  |
| Db | 1050 | TTCGCCAGCGTCCTTGGAAATCTGCCAAATATTTCTTACTAGCTCAACACCTCTTAT        | 1109 |
| Qy | 961  | ggaccttggagaagtttctaagtttgaatctctcccaacaacatgataagctatga         | 1020 |
| Db | 1110 | GGACCTTGGACAGTTTCTACTGTTTTTGTCTATCTCTTCCACACAACTGATGCTATGSA      | 1165 |
| Qy | 1021 | agctatgtlcaaaatagaacagctgtcccaaggagaacccaactaagatlgaaaaataatga   | 1080 |
| Db | 1170 | AGCTTATGTCAAAAGTAGACAGCTGTCCAGAGGAACCCCAACTGCAATGAAAAATAATGA     | 1229 |
| Qy | 1081 | agaagcggaaagactatgaatgatacttaactgaattcttgaatgagatgtygttaagtttga  | 1140 |
| Db | 1230 | AGAAACCGGAACATATGATATGATGATCTTACTGATTTCTGAAATGAGATGTGTCAGTTTGA   | 1289 |
| Qy | 1141 | tgaatgaacactctctcccttcatccaacttgcataagtttgccaagaagaactccctaaac   | 1200 |
| Db | 1290 | TGATGACACTCTCTCTTCTTTATCCAAATTTCCCTAGTTGCCAAGACATCTCTTAAC        | 1349 |
| Qy | 1201 | ttgggttaeatatactgtcgtgaagaaggagactgctgactatgctccctagtcctgc       | 1260 |
| Db | 1350 | TTGGGTACATTTACATTTGCTGCTGTAAGAGGAGACTGGGACTATMGCTCCCTTAGTCTGCG   | 1409 |
| Qy | 1261 | ccccgaatgacgaagattataaagltcaataattgaacaatgagccctcaagcggatgtgtag  | 1320 |
| Db | 1410 | CCCCGATGCGAGAGTTATATAAAGTCAATTTTGAACAATATGGCCCTCAGCGATGTGTAG     | 1469 |
| Qy | 1321 | gaatatacaaaaatctcgattttagatatacaagaatgaacactttaagactcgtgaagc     | 1380 |
| Db | 1470 | GAAAGTCAAAAAATCCGATTTATAGGATATACAGATGAAACCTTTAAGACTCGTAGAGC      | 1529 |

QY 1381 taccagatcaatcaagaatcttggaccttaacttaattgggaagtggagacact 1440  
|||||  
Db 1530 TATTACAGATCAATCAGGAATCTTGGACCTTTATGGAAGATTGGAGACACT 1589  
QY 1441 gtgatataatcaagaatcaagaacagacatacaactcaactcaggaatcac 1500  
1590 GTGATATATATTAAGATCAAGACAGACATATACCTACCGATCAC 1649  
QY 1501 tgatgtccgtcttctgtaatcaagaagattacaaaagtgtaaaacattgaagattc 1560  
1550 TGATGTCCGTCCTTGTATCAAGAGATTACAAAAGGTGTAAACATTTGAAGGATTT 1709  
QY 1561 tccaatctgcagagagaatattcaataataatgagacagtgactgaagaatgggccc 1620  
1710 TCCAAATCTGCAGAGAAATATTCAATATATATGACATGACTGTAGAGATGGGCC 1769  
QY 1621 aactaaatcagatccctgctgctgacccgtataacttaactgcttcaatgaagag 1680  
1770 AACTAAATCAGATCTCGTGCTGACCCGCTATTACTGATTGCTTATATGAGAG 1829  
QY 1681 agatctagctcaggaactcattggccctctccatctgtctcaagaatctgtatga 1740  
1830 AGATCTAGCTTCAGAGACTCATTTGGCCCTCTCCATCTGCTACAAAGATCTGTAGATCA 1889  
QY 1741 aagaagaaacagagataatgtaagaagaagaatgtaacatctgtttctgtattatga 1800  
1890 AAGAGAAACAGATATATGTCAGACAAAGAGATGTCACTGTGTTCTGTATTTGATGA 1949  
QY 1801 gaacgaagctgtaacctcaagaaatatacaagcttctccccaatccagcttgagat 1860  
1950 GAACCGAAGCTGTAACCTTCACAGAGATATACAGCGTTCTCCCAATCCAGCTGGAGT 2009  
QY 1861 gcagcttgagagatccagaagttccaagcctccaacatcatgcaagcatcaatgagat 1920  
2010 GCACCTTGAGAGATCCAGAGTTCAGAGCCTCCACATCATGACAGCATCAATGCTATGT 2069  
QY 1921 ttgtgagattgagctgtaagttgtttgacagagtgagagacatgtaacttcaatga 1980  
2070 TTTTGTATGATTTGCATGTTGTCAGTTGTTGTCATGAGTGGCATATCTGTATCTTAAG 2129  
QY 1981 cattgagacagagactccttcttcttcttcttcttcttcttcttcttcttcttctt 2040  
2130 CATGAGAGACAGACTGATCTCTTCTGTTCTTCTGTCATGATATACCTTCAACAA 2189  
QY 2041 aatgcttataagaagacacactcaacctatcccatcttccagagaaactgtctcattgc 2100  
2190 AATGCTATGAAAGACACACTCACCTATTCCTCATCTCAAGAGAAACTGTCTCATGTCT 2249  
QY 2101 gatgaaacccaggtctatgagattctgaggtgagacacactcagacttctggaacagag 2160  
2250 GATGGAACCCAGGCTATGATGATTTGAGGTTGACCAACTCAGACTTTCGGAACAGAG 2309  
QY 2161 catgaccgcttactgaaggtttctagttgtgacaagaacactgtaattatlaacagaga 2220  
2210 CATGACCGCTTACTGTAAGGTTTCTAGTTGTGACAAGAACACTGCTATATATTCGAGGA 2369  
QY 2221 caattatagaagataattcagacatactgtctgagtaaaacaatgcaattgaacaaag 2280  
2270 CAGTTATGAAGATATTTACACATATCTGTGAGTAAACAAATGCCATTGAAACCAAGAG 2429  
QY 2281 ctctcccaagaatcaagacacccctagacactaggaacaaagaatlaattccaccca 2338  
2430 CTCTCCCAAGATTCAGACACCTTAGCCTAGGCAAAAGCAATTTATGTCACACACA 2487  
Db

RESULT 13  
LOCUS I31901 9009 bp DNA PAT 06-FEB-1997  
DEFINITION Sequence 3 from patent US 5583209.  
ACCESSION I31901  
VERSION I31901.1 GI:1822692  
KEYWORDS

SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 9009)  
AUTHORS Lollar, J.S. and Runge, M.S.  
TITLE Hybrid human/porcine factor VII  
JOURNAL Patent: US 5583209-A 3 10-DEC-1996;  
FEATURES  
source location/qualifiers  
1..9009  
BASE COUNT 2853 a 1907 c 1844 g 2405 t  
ORIGIN  
Query Match 52.9%; Score 2336.4; DB 10; Length 9009;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2337; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 catgcaatagagctctccacgtcttcttcttcttcttcttcttcttcttcttcttcttctt 60  
150 CATGCAATTAAGAGCTCTCCACCTGCTTCTTCTGTGCTTGTGCGATTCTGCTTTAGTGC 209  
QY 61 caccagaagatactacctgagtgagtggaactgcaatggaactatataatgaagtgatct 120  
210 CACCAGAAGATACCTAGCTGGGTGCAAGTGAAGTGCATGAGATATATGAAAGATGATCT 269  
QY 121 cgtgtgagctgctgctgagagcaagatcttctccttagagtgccaaatcttccatcaa 180  
270 CGGTGAGCTGCTGTGAGAGCAAGATTTCTCTCAGAGTGCCTCAAAATCTTTTCCATTCAA 329  
QY 181 caactcagctgctgtaacaaagactctgtttagaagatcaagcttcaacttcaacat 240  
330 CACTCAGCTGTGACAAAAGACTCTGTTGTGAAATTCACGGTTACCTTTTCAACAT 389  
QY 241 cgttaagcacaagcgcacctgagatgagctgctagagcttcaacatccagagctgagatla 300  
390 CGTTAAGCAAGGCGACCGCTGATGGGTGCTGATGCTTACATCCAGCTGAGGTTTA 449  
QY 301 tgataagtgcttacttaacttaagaacatgagcttcccatctctgtcagcttcaatgctgc 360  
450 TGATTAAGTGTGATTAACATTAAGAACATGAGCTTCCATCTCTGCACTTTCATATGCTGT 509  
QY 361 ttgttcttccactggaagactctgagagagctggaatgatgatgaacagctcaag 420  
510 TGTGCTATCTCTACTGGAAGCTTCTGAGGAGCTGGAATATATATATGACAGACTCAAG 569  
QY 421 ggaagaagaagaatgaatgcttccctgctgagagcacaatataatgcttgcagagctc 480  
570 GGAGAAAGAAAGATGATTAAGCTTCCCTGAGGAGCATATATGCTGCGAGGTCT 629  
QY 481 gaaagaagaatggtccaatgagcctctgaacccactgtgccttaactcaatcttcttca 540  
630 GAAAGAAATGTCATATGAGCTTGCACCTGACCTGCTTCTGCTTCTGCTTCTTCTCA 689  
QY 541 ttgtgagctgctgtaagaagactggaatcagagcctcaattgagcccttaactagatlaaga 600  
690 TGTGAGCTGCTGTAAGAACTTGAATTCAGGCTTCAATTTGAGACCTACTAGATATGAGGA 749  
QY 601 aaggagcttgagcacaaggaagaacacagacttgacaaatlaactacttcttctgtc 660  
750 AGGAGTCTGGCCTGCAAGGAAAGACACAGACTTGCACAAATTTATACACTTTTGGCTGT 809  
QY 661 atttgatgaagggaaagtgtggaactcaagaacaaagaactccttgaatgaggaatagga 720  
810 ATTTGATGAAGGAAAGGTTGGCACTCAGAAACAAAGAACTCTGTGATGCGAGATAGGGA 869  
QY 721 tgcgtcatctgctgagcctgagcctaaatgcaacagctcaatgtaatgtaagaacagctc 780  
870 TGTGCTATCTGCTGGGCTGAGCTTAAATATGACACAGCTCAATGCTATATGTAACAGTCT 929  
QY 781 tctgccaagcttgaatgagatgcaagaagaatcagctcattgagcatgtaattgaaatgg 840  
930 TCTGCCAGGTCTGATGATGCCACAGGAAATCAGTCTATGTCATGTGATTTGGAATGGG 989  
Db

|    |      |  |      |
|----|------|--|------|
| QY | 841  | caccactctgtgaagctgcactcoaalatccctcgaaagtlcaacatlctltgtgaagacc    | 900  |
| Db | 990  | CACCACCTCCTGAAGTGCACTCAAAATTCCTCGAAGGTCACACATTTCTGTGAGAACCA      | 1049 |
| QY | 901  | tgcgaagggctcttgaaabctgcgaataacttctctactctgctcaaaacttgat          | 960  |
| Db | 1050 | TCGCCAGGCTCCTTGGAAATCTGCCCAATTACTTTCCTACTGCTCAACACTCTTAT         | 1109 |
| QY | 961  | ggaacttggaacglttctactglttltgcalatctctcccaacaatga tggcatlga       | 1020 |
| Db | 1110 | GGACCTTGGACAGTTTCTACTGTTTGTGCATATCTCTCCACCAACATGATGGCATGA        | 1166 |
| QY | 1021 | agcttatgtlcaagtagaacgctgtgccagagaaacccaactaagaa tgaataataga      | 1080 |
| Db | 1170 | AGCTTATGTCAAAAGTAAACAGCTGTCCAGAGAACCCCACTACGAAATGAATAAGA         | 1229 |
| QY | 1081 | agaaggggaagactatgatgatgtctactgattctgnaatgtgatgtgtgcagttlga       | 1140 |
| Db | 1230 | AGAGCGGAGACTATGATGATGATCTTACTGATTCGAAATGGAATGTGTGACAGTTTGA       | 1288 |
| QY | 1141 | tga tgaacaactctctctctcttcaacaaatgcgtcagltgtccaagaaacatccta       | 1200 |
| Db | 1290 | TGATGACAACTCTCTCTCTCTTATCCAAATTCGCTCAGTTGCCAAGAAAGCTCTAAAC       | 1349 |
| QY | 1201 | ttaggtatcaatcaatgtctgtctgaagggggaatggagactatgctccctagtctctgc     | 1266 |
| Db | 1350 | TTGGGTACATTACATTTGCTCTGTGAAGGAGAGACTGGGACTATCTCTCTTATGTCTGC      | 1409 |
| QY | 1261 | ccccga tgaacaagatataaaatcaatatltbaacaatgcccctcagcgatgtgtag       | 1320 |
| Db | 1410 | CCCCGATGACAGAAAGTTATTAAGACTCATATTTTAACATATGGCCCTCAGCGATTGGAG     | 1469 |
| QY | 1321 | gaagtaacaaaaagctccgattatltgcatcacacagatgnaaccccttaagactctgtgaagc | 1380 |
| Db | 1470 | GAAGTACAAAAAGTCCGATTTATGCGATACACAGATGAACCTTTAAGACTCGGAGAGC       | 1529 |
| QY | 1381 | tatcagcatgaatcaggaatccttgggaacttcaactttaggggaagtltgaagacaact     | 1440 |
| Db | 1530 | TATTTGACCATGAATCAAGAAATCTTGGGACCTTTCCTTATGCGGAAGTTGGAGACACT      | 1589 |
| QY | 1441 | gtgtatataatlaagaatcaagaacagacagacatatacaatcaccctcaggaatcac       | 1500 |
| Db | 1590 | GTTGATTATATTTTAAGATCAAGAACACAGACCATATTAACATCTAACCTCACGGATATAC    | 1649 |
| QY | 1501 | tga tgcgcgtccttctgtatccaaggagatataccaaaagltgaataacatttgaagatct   | 1560 |
| Db | 1650 | TGATGTCCGCTCTTGTGATTTCAAGGAATTAACAAAAAGGTGTAATAACATTTGAAGGATTT   | 1709 |
| QY | 1561 | tccaatcttcgcaggagaaatattcaaatlaaa tgaacagtgaacttaagaaga tggacc   | 1620 |
| Db | 1710 | TCCAAATTCGCCAGGAATAATTTCAATTAATTAATGAGACGTGACTGAGAGATGGCC        | 1769 |
| QY | 1621 | aactaaatcagaatccctcgtgtgcctgaaccgctatctactctagtcttlaa tgaagag    | 1680 |
| Db | 1770 | AACTAATATGAGATCCTCGGTGCGTGAACCCGCTATTACTGATGTTTCGTTATATGAGAG     | 1829 |
| QY | 1681 | agatcagcttcaagactcatgtgcctctcccaatctctgtcacaagaatctgtatagatca    | 1740 |
| Db | 1830 | AGATCTAGCTTCAGGACTCATTTTGCCCTCTCTCATCTGTCTACAAABANCTGTGATGATCA   | 1889 |
| QY | 1741 | aagagaagaaccagaataatgtcagaagaagagaatgatactctgttcttctgatttltga    | 1800 |
| Db | 1890 | AAGAGAAACCAATATATGATCAACAGAGGAATGTATCTCTGTTTCTGTATTGTATGA        | 1949 |
| QY | 1801 | gaaccgaagctgttactcctcagagaaatacaacgcttcttcccaatccagcttgagt       | 1860 |
| Db | 1950 | GAAACCAACTGTGTACTCAACAGATATATAACAAGTTTCTGCCCAATCAGCTGAGGT        | 2009 |
| QY | 1861 | gcagcttgaagatccagagttccaaagcttcaacatcatgcaacagcatlcaatgtgctatgt  | 1920 |
| Db | 2010 | GCAGCTTGAGATCCAGAGTTCCAAGCTCTCAACATCATGATCAGACAGCATTAATGCTATGT   | 2069 |

|                            |                                     |  |                     |
|----------------------------|-------------------------------------|--|---------------------|
| Qy                         | 1921                                | ttttgtaagtgtgcagttgtgtaagttgtttgttgtaaggggtgagtaactggttaacttctaa | 1980                |
| Db                         | 2070                                | ttttgattatgtttgcagttctcagttgtgttgatgagtggtggtatctggtatattctttaa  | 2129                |
| Qy                         | 1981                                | catgtgagcaacagactgacttctctctgtctctctctctctgtatataacttcaaacaca    | 2040                |
| Db                         | 2130                                | catgtgagcaacagactgacttctctctctctctctctctctctgtatataacttcaaacaca  | 2189                |
| Qy                         | 2041                                | aatgtgtctatgaagacacactacccctattcccatctctcagaagaaactgttctcatgtc   | 2100                |
| Db                         | 2190                                | aatgtgctrtatgaagacacactacccctattcccatctctcagaagaaactgttctcatgtc  | 2249                |
| Qy                         | 2101                                | gatgtgaaaccccaagtgctctatgtatcttgggtgtggccaaactaagactcttcgaaacaag | 2160                |
| Db                         | 2250                                | gatgtgaaaccccaagtgctctatgtatcttgggtgtggccaaactaagactcttcgaaacaag | 2309                |
| Qy                         | 2161                                | catgacagcgtctacttaagaagttctctagtgtgtacaaagacactgtgtatattatcagaga | 2220                |
| Db                         | 2310                                | catgacagcgtctacttaagaagttctctagtgtgtacaaagacactgtgtatattatcagaga | 2369                |
| Qy                         | 2221                                | cagttatgagaalatttcagacatactgtctgtgaataaacaalgtccattgaaccaagaag   | 2280                |
| Db                         | 2370                                | cagttatgagaalatttcagacatactgtctgtgaataaacaalgtccattgaaccaagaag   | 2429                |
| Qy                         | 2281                                | ctttcccccgaattcaagaacaccctcagactcctgtggcaaaagcaatttaagccaccaca   | 2338                |
| Db                         | 2430                                | ctttctcccagaaattcaagaacaccctcagactcctgtggcaaaagcaatttaagccaccaca | 2487                |
| RESULT                     | 14                                  |  |                     |
| LOCUS                      | 163424                              | 9009 bp  | DNA                 |
| DEFINITION                 | Sequence 1 from patent US 5663060.  |  |                     |
| ACCESSION                  | 163424                              |  |                     |
| VERSION                    | 163424.1                            | GI:2480997   |                     |
| KEYWORDS                   |                                     |  |                     |
| SOURCE                     | Unknown.                            |  |                     |
| ORGANISM                   | Unknown.                            |  |                     |
| REFERENCE                  | 1 (bases 1 to 9009)                 |  |                     |
| AUTHORS                    | Lollar,J.S. and Runge,M.S.          |  |                     |
| TITLE                      | Hybrid human/animal factor VIII     |  |                     |
| JOURNAL                    | Patent: US 5663060-A 1 02-SEP-1997; |  |                     |
| FEATURES                   | Location/Qualifiers                 |  |                     |
| source                     | 1..9009                             |  |                     |
| BASE COUNT                 | 2853 a                              | 1907 c   | 1844 g              |
| ORIGIN                     |                                     | 2405 t   |                     |
| Query Match                | 52.9%;                              | Score 2336.4;  | DB 10; Length 9009; |
| Best Local Similarity      | 100.0%;                             | Pred. No. 0;   |                     |
| Matches 2337; Conservative | 0;                                  | Mismatches 1;  | Indels 0; Gaps 0;   |



|    |      |  |      |
|----|------|--|------|
| Qy | 301  | tgtatagatggtcattacttaacttaagaacatgcttcccatcctctgacgtcttcaatgct     | 360  |
| Db | 450  | TGATACAGGTGCTATTACACTTAAAGACATGGCTTCCACTCTGACAGTCTTACGTGT          | 509  |
| Qy | 361  | tgtgtatcttacttggaaaagctcttgaaggagctgaatatgatactgaacacagtcgaag      | 420  |
| Db | 510  | TGtGTATCCtACTGGAaACCTTCTGaGGAGAGCTGaTATGaTGTaGATCaGACaGTCaAG       | 569  |
| Qy | 421  | ggagaaagaagatgataaagctctccctggctgggaagccaatactctctggcaggtcct       | 480  |
| Db | 570  | GGaAAAGaAGATGaTAAATCTTCCCTGGTGGAAACCaTACATATGTCTGGAGAGTCT          | 629  |
| Qy | 481  | gaaggaatctgtccaaatggccctctgacccacatggcccttacctatacttcttcca         | 540  |
| Db | 630  | GAAGAGAAATGTGTCCAAATGAGCCCTCTaCCACTGTGCTTACTTACTATCTTTCTTCA        | 689  |
| Qy | 541  | tgtggaccttggtaaaagacttgaatcaggtccatctggagccctactagatctgtaga        | 600  |
| Db | 690  | TGTGTGACCTGTGTAAAGACTTGAATTCAAGGCTCATTTGAGGCGCTTACTAGTATGAGAGA     | 749  |
| Qy | 601  | agggagctctggccaaggaaaagacaagaccttgaccaaattatactacttctgtct          | 660  |
| Db | 750  | AGGGAGTCTGGCCAGAGAAAGACACAGACTTGGCAAAATTTATACCTTTTGTGCTGT          | 809  |
| Qy | 661  | atttgaagaaggaaaaagtctggcacccagaaaaagaacatcctctgaatgacagatagga      | 720  |
| Db | 810  | ATTTTGATGAAGGGGAAAAGTTGGGCATCAGAAAACAAAGAACTCTTGATGACAGATAGGGA     | 869  |
| Qy | 721  | tgtctgcatctgtctcggcgcttgctctaaatatgcacacagtcacatgtgtatgtaaacagctc  | 780  |
| Db | 870  | TGCTGCATCTGCCTGGGGCCCTGGCTTAAATGACACAGTCATGGTTATGTAAACAGGTC        | 929  |
| Qy | 781  | tctctcagagctctatctggatgacccaggaacatcagtcatactgcatctggatctggaaatggg | 840  |
| Db | 930  | TCTGTCCAGGTCTATTGTGATGGCCACAGAAATCAGTCTATTGGCATGTGATTTGGAAATGG     | 989  |
| Qy | 841  | caccactctgtaagctgcaactcaatatctctcgaaggtcacaacatctctgttgaagaaca     | 900  |
| Db | 990  | CACCACCTCTGAAGTGCACACTCAATATTCCTCGAAGTCAACATTTCTGTGAGAGAACCA       | 1049 |
| Qy | 901  | tcgcgaaggctcttggaaatccgcgcaataactctcttacttgcttgaacaaccttgat        | 960  |
| Db | 1050 | TCGCGAAGGCTCTTGGAAATCTCGCCAAATCTTTCTTACTGCTCAACACTCTTGAT           | 1109 |
| Qy | 961  | ggaccttggacagcttctactgctgttctgtgataatactcttccaccaacaatgataggcatgga | 1020 |
| Db | 1110 | GGACCTTGGACAGTTCCTACTGTGTTTGTCAATATCTTCCACCACCAATGATGGGATGGA       | 1169 |
| Qy | 1021 | agcttattgtcaaaagttagacagctgcttcagaagaaacccaactagaatgaaaataatga     | 1080 |
| Db | 1170 | AGCTTATGTCAAAAGTAGACAGCTGTCCAGAGAACCCCAACTACCAATGAAGAAAATATGA      | 1229 |
| Qy | 1081 | agaagcggaagactatgatgatgatcttactgactctcgtagaatgagatgtgtcgaatttga    | 1140 |
| Db | 1230 | AGAGCGGAGACTATGATGATGATGATCTTACTGATCTTGAAATGAGATGTGTGTCAGGTTTGA    | 1289 |
| Qy | 1141 | tgatgatcaactctctctccttcttccaaatctcgctcgaattgtgcgaagaagacatcctaaac  | 1200 |
| Db | 1290 | TGATGACACACTCTCTCTCTTATCTCCAAATTTGGCTCAGTTGGCAAGAGACATCTCTTAAAC    | 1349 |
| Qy | 1201 | tttggatcattatcatctgtctgtcgaagaagaagacatgggaactatgtctcctaagtccctgcg | 1260 |
| Db | 1350 | TTTGGGTACATTACATTGTCTGTGAGAGAGAGACTGTGGACTATGTCTCCCTTAGTCTCTGC     | 1409 |
| Qy | 1261 | ccccgatgacagaagattataaaagtcaatcatatttggacaatgtgcccctcagcggatcttgg  | 1320 |
| Db | 1410 | CCCCGATGACAGAGATTATAAAGTCAATATTTGGAACATGTGCCTCAGCGGATTTGGTGA       | 1469 |
| Qy | 1321 | gaatatcaaaaaagtcogatttttttggctatcacagaatgaatgaactttaagactcgtgaagc  | 1380 |
| Db | 1470 | GAAATCAAAAAAGTCCGATTTTATGGCATACAGATGAACCTTTAAGACTCGTGAAGC          | 1529 |

|    |      |  |                                 |      |
|----|------|--|---------------------------------|------|
| Oy | 1381 | taacacatgaataacaggaatctctggacccttactctta   | tgaggagaatttgagaaacact          | 1440 |
| Db | 1530 | TATTTCAGCATGAATACAGGAATCTTTGGGACCTTTACTTTT | TGGGGAGATTGGAGACACT             | 1589 |
| Oy | 1441 | gtgatataatcttaagaataaagaacagacacata        | taacatctacacctacagatatac        | 1500 |
| Db | 1590 | GTTGATTATATTTAAGAATCAAGCAAGACCATATTA       | CATCTACCTACCGGAATCAC            | 1649 |
| Oy | 1501 | tgatgctccgtcccttgatctcaagaagatacc          | aaaggctgaanaacatltgaagaatt      | 1560 |
| Db | 1650 | TGAATGCTCCGTCCTTTGTATTCAAGGAGATTAC         | CAAAAGGCTTAACAACTTTGAAAGGATTT   | 1709 |
| Oy | 1561 | tccgaattccccaagaaaataatctaaataataatga      | atgaatgaatgaagaatggacc          | 1620 |
| Db | 1710 | TCCAATTTCTGCACAGACAATTAATTTCAATTAAT        | TGACAGATGTPAAGATGGGCTC          | 1769 |
| Oy | 1621 | aactcaatcagatccctcggtgacctgaacccgcta       | atctactagtcttcgttaatatgaag      | 1680 |
| Db | 1770 | AACTAATTCAGATTCCTCGGTGGCTGACCCGCTAT        | TACTTCTAGTTTCGTTAATATGGAAG      | 1829 |
| Oy | 1661 | agatcagctcagagacatctgcccctccctcaatc        | tgctacaaagaatctgtagatca         | 1740 |
| Db | 1830 | AGATCTAGCTTCAGAGACTCATTTGGCCCTCTCC         | TCATCTGCTCAAAAGATCTGTGTAGTCA    | 1889 |
| Oy | 1741 | aagaggaaacagataatctgcacaagaaggaaatg        | catccctggtttctgctatttgata       | 1800 |
| Db | 1890 | AAGAGGAACCCAGTAATGTGTACAACAAGGAAT          | GTCACTCGTTTCTGTATTTTGAGTA       | 1949 |
| Oy | 1801 | gaaccgaagctggtacctcaacagaaatatacaag        | ctcttcctcccaatccagctgagat       | 1860 |
| Db | 1950 | GAACCGAAGCTGTACTCTCAACAGAAATATAC           | AGACGCTTTCGCCATTCAGCTGAGCT      | 2009 |
| Oy | 1861 | gcagcttgaagatccagagttccaaagcctccaac        | atcalgtacagacatcatatgctatgt     | 1920 |
| Db | 2010 | GCACCTTGAGAGATCCAGAGTTCCAAAGCCTTCC         | AAATCATCATGACACAGATCAAGGCTATGT  | 2069 |
| Oy | 1921 | tttctgaatgtttgaagtctgcagtgtgtttgtga        | ctgagggtggagatactgatacttaag     | 1980 |
| Db | 2070 | TTTTGATATGTTTGAGATTGTGAGTTGTGTTGTG         | CAATGAGGTGATCTGTATCTATTTAAG     | 2129 |
| Oy | 1981 | cattgagacacagactgaactccctctctgtctctc       | ctctctctctctgtgataaccttcaaacaca | 2040 |
| Db | 2130 | CATTGGAGCACAGACTGATCTCCCTTCTGTCTT          | CTCTCTGTGGATATACCTTCAAAACATA    | 2189 |
| Oy | 2041 | aatggtctatagaagaacacactacacctatctcca       | ttctctaagagaatctgtctcaatgtc     | 2100 |
| Db | 2190 | AATGCTGTATGTAAMACACTCTACCTATTCCTAT         | TCTCAGAGAAACGTCTTCATGTCTC       | 2249 |
| Oy | 2101 | gatggaaaaccccggtctatagaattctgggtg          | ggcccaactaagaactttcggaanaaggg   | 2160 |
| Db | 2250 | GATGGAAAAACCCAGCTTATAGGATTTCTGGGG          | GTCCCAACTAGACTTTTGGAAACAGAGG    | 2309 |
| Oy | 2161 | catgacgcaccttactgaaggctctctagttgtga        | caagaacactgtgtatattacgaaga      | 2220 |
| Db | 2310 | CATGACCGCCTTACTGTAAGGTTTCTAGTTGTG          | ACAAGAACACTGGTGATTATACGAGTA     | 2369 |
| Oy | 2221 | cagttatgaagaataattcagcatactctgctgag        | ttaaanaacaaatgcatltgaaccaagag   | 2280 |
| Db | 2370 | CAGTTATGAAGATATTTTCAACATACTTTCG            | TGAGTAAACATATGCCATTTGAACCAAGAG  | 2429 |
| Oy | 2281 | ctctctccagaattcaagaacaccttgcaactg          | ggcaaaagaatltatgcaaccaca        | 2338 |
| Db | 2430 | CTTCTTCCAGAAATTCAGACACCTTGAACCT            | TAGGCAAAAGCAATTTAATGCCCCACCA    | 2487 |

| RESULT 15  |   |           |             |
|------------|---|-----------|-------------|
| LOCUS      | HOMEVIIII   |           |             |
| DEFINITION | Human coagulation factor VII:C (anti-hemophilic factor) | 9009 bp   | 08-NOV-1994 |
| ACCESSION  | K01740  |           |             |
| VERSIONS   | K01740.1  | GI:182802 |             |
| KEYWORDS   | factor VIIII; hemophilic.                               |           |             |
| SOURCE     | Human cDNA to liver mRNA, and genomic DNA.              |           |             |



| ORGANISM        | Human sapiens  |
|-----------------|--|
| REFERENCE       | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 9009)  |
| AUTHORS         | Toole,J.J., Knop,J.J.L., Mooney,J.M., Sultzman,L.A., Buecker,J.L., Pletman,D.D., Kaufman,R.J., Brown,E., Shoemaker,C., Orr,E.C., Amphlett,G.W., Foster,W.B., Coe,M.L., Knutsen,G.J., Fass,D.N. and Hewick,R.M.   |
| TITLE           | Molecular cloning of a cDNA encoding human antithaemophilic factor   |
| JOURNAL MEDLINE | Nature 313 (1984), 342-347 (1984)  |
| COMMENT         | Currently hemophilias are treated with factor VIII:C purified from blood. The human factor VIII:C cDNA has now been isolated, cloned and expressed. The gene was also isolated. It consists of a large number of exons spanning over 180 kb. The authors kindly supplied the sequence directly to Genbank by computer network.   |
| FEATURES        | Location/Qualifiers  |
| source          | 1..9009  |
| gene            | /organism="Homo sapiens"   |
| CDS             | /db_xref="taxon:9606"  |
|                 | /map="xq28"  |
|                 | 151..7206  |
|                 | /gene="F8C"  |
|                 | 151..7206  |
|                 | /gene="F8C"  |
|                 | /note="Factor VIII"  |
|                 | /codon_start=1   |
|                 | /db_xref="db:G00:119-124"  |
|                 | /protein_id="AA52484.1"  |
|                 | /db_xref="GI:182803"   |
|                 | /translation="MQIEELSCFLLCLRCFSATRYLYGALVLSMDYMQSDLGELPVDAREPRVRSPEPNTSVYKTLLEVEFVHLHNIAKPRPMMGLGPTQAEVDTVTITLKNSHPVSLHAVGVSVKWSGAEIGDTSOREDEKVPFGSGHTYVMQLKENPMASDPLCLTYSLSHVDIAKDNSGELIGLALCREGSLAKETQTLHFILFAVEFGKSMHSEKNSLMODDASASAMPMPVNCVYNSLDELGLGPKRSHYMHVIGMETPEVHSLIELEGHDFLVNRHRSALSTSPTEPTLAOTLMDGCLPILKISHOHDMEMAYVKVDSCPEERPOLRMKNEAEYDDDDITSEMDYVAFDDDNPSFIOTRSVAKHNPVTHVIAAEEEDWYAPLVLAIPDRSTSOYLNNGPQIRCKTKVREPMVTDEFTKREALYQHSIGILGLLYGEVDTLIFKNQASRYNPIPGIDVPLRFLRPGVHLKDFPILPEIFEKYKWTVAVEGDPTSDRCRLTRYKSSVNMREIASGLIGPLTICYKEVDROGNOIMSDKRNVLISVFEDNRSMVLENTORFLPNPAQOLDEPEPASNIMHSINGVFDLSLOVCLHEAVYMTVLTLSGAOTDPLTSFEGSTPKKKHYVEDTLFPPSGEYVPMSEMRPGMLTIGCHNSDPRNKGATLAKVSSCDKNTGYDEDSEYDLSVYLSKNNALIEPRSSONSRRPSTROKQFNAVTTPENDTEKTDWEAFRRPMPKIDNVSSSDMLMLRQSPPHGLSLSDLOEAKETSDPSPGAIDSNLSSEMHFRLPQHSNGDWLTPEESLOALRLEKTLGTATLETKKDFPVSSNNLITIPSDNLAAGTNTSSSLGPPSMVHYDSOLDOTTLTGKSSPLTESGGLTSEKNDKLESGIMNSOSSMGKTVSPESGRLEFKGRAGHPALLTDKMLPKVYSILKTKNKSNNSATRKTHIDGSLIENSPSYQNIULESDTEFKVTPILIDRMLMDKNAITLRINHNSKTTSSKMEWYQOKKEGPIPPDQNPDMPSFFKMLPEPSARWLQTHGKNSLNSGGPSRQVLSLGEKSPSEGNFLSEKKNVYGGEFTKGLKEMVFPSSRNLNINLNLHENNTNHOEKRIQIELEKETLLOENVLQITVGTCKFMKMLFLSTRONVESYBGAAYPLDOPRSLINDSTNRKTKHTAHFSKKGEEENEGLOKTOIVEKVACTRISPTSCQNFVQRSKRAKQPRLEPELELEKRTIIVDSTQMSKNNKHLPTSLTODIYNEKEKATIQSPLESDCLTSHSTIPQANNSPLPAKVSFSPSTIRPLTVLVEQDMSHPLAASATRRKDSGVQESHPLQGAQRNNLSIALITLMTEDDQEVSLGTSANSTYKVKENLVLPKDLPTKTSGLKVELPKVHIYKQDLPETSGSGIDELVGLSLQGTAKIKNANLPKAGKVPFLRVATIESAKPSPLDLAMDNIYGNQIPREKSOEKAPRETAIPKRTDLSLNAGCSNHAIAINCGOKPRLVYIAQOGRTERLCSNPVLAKHOREITRTTLDSDDEIDYDITISYEMKEEDPDTIDEDENOSPFSOKRHTHTAIVERLMDYGNSSSPHVLNRNAGSGVQPEKRVVQFEDTSFQPLYRELNELNLADEPLYIRA EVDENIVATFRNQSAPRSYFSSLSLIEEDQROGAERKNVFPVNEKTYFKVQVNLAPTEDEDCAMAEFSDVLEKDVHSLIPDLVCHNTLNPHAGROVTEQALFPTIIPDTSKSWYFLENMERNCRAPCNIOMTQVEMPRHALINQYIMDTLPCLMAAORIRKWLKSMYENENHSTHPSGHTVYKKKEKYNALNLPVPEYVLELPSAKATIRVECLIGHLHAGMSTLFLVYSNKCQDPLGMAISGIRBOFQTAGSGQVAMPKLARLHYSGSISNASTKEPSWIKVLDLAIPMIINGITQGAQKQFSSLYISOFTINYSLDGKQW TYRNSNATGLMVEFGANDSGIKHNIPIGTIRYILALHTYHSISRTLEMLMGCDL NSCGMPLGMSKSAISDAOTIASSFTYMTKFPWSSKPLHLDQSNANRPQVNNPKEMLQVDFOKFMTKVTGTTQGVKSILTSMVKKRFLISSSDGQHMPLFQNGKVKYKFGQNDSPRYVNSLDPLTRYLALRHQSWYHQAIALKNEVLCGAQDLY" |
| BASE COUNT      | 2853 a 1907 c 1844 g 2405 t  |
| ORIGIN          | 160 bp 5' to 3' for SacI cut site; chromosome Xq28.  |

[illegible]

QY 1021 agcttatgcaaaagtagacaagctgtccagagaaacccaactacgaatgaaaaataatga 1080  
|||||  
Db 1170 AGCTTATGTCTCAAACTAGACACTCTCCAGAGAACCCCACTACGAATGAAAAATTAATGA 1229  
QY 1081 agaagcggaaagactatgatgatagtcttactgaltctgaaatggaatggttcaggtttga 1140  
|||||  
Db 1230 AGAAGCGGAAGACTATGATGATGATCTTACTGATTTCTGAATGATGATGGTCAGGTTTGA 1289  
QY 1141 tgatgacaactctctcccttccatccaaatcgctcagttggccaagaagaactcctaaac 1200  
|||||  
Db 1290 TGAGAGAACACTCTCTCTTATCCAAATTCGCTCAGTTGCCAAGAAGCATCTCTAAAC 1349  
QY 1201 ttgggtacatcacatgtctgtgaagaaggaggaactgggaactgctcccttaagctccgcgc 1260  
|||||  
Db 1350 TTGGGTATCATTTACATTCCTGCTGTGAAGAGAGACTGGGACTATCTCTCTTAAGCTCTGCC 1409  
QY 1261 ccccgatgacagaagatataaagtcacatattgaaacaatggccctcagcggatltggtag 1320  
|||||  
Db 1410 CCCCAGATGACAGAGATTATTAAGTCAATATTTGAACAATGGCCCTCAGCGGATTTGGTAG 1469  
QY 1321 gaagtacaaaaaagctccgattttagtcatacacagatgaaacctttaagactgtgaagc 1380  
|||||  
Db 1470 GAAGTACAAAAAAGTCCGATTTATGCTACACAGATGAAACCTTTAAGACTCGTGAAGC 1529  
QY 1381 tattcagcatgaaatcaggaatcttggaccttacttlatggggaagttgagacacact 1440  
|||||  
Db 1530 TATTACACATGAATCAGAGATCTTGGGACTTTTATGGSAASTTGGAGACACACT 1589  
QY 1441 gtgatataatlttaagaatcaagcaagcagacacataaacatctaccctcacggaalcac 1500  
|||||  
Db 1590 GTTGATTTATATTTAAGATCAAGCAAGCAGACCATATATACATCTACCCCTCAGGAATCAC 1649  
QY 1501 tgatgtccgtcccttggatcacaaggagatacccaaaagtgtlaaacaatttgaagaatc 1560  
|||||  
Db 1650 TGAGTCCGTCCTTTGATTTCAAGGAGATTACCAAAAGGTAAACATTTGAAGGATTT 1709  
QY 1561 tccaatctcgcagagaagaatatcaaatataaagtgaactgtatagaatltgggcc 1620  
|||||  
Db 1710 TCCAAATCTGCCAGAGAAATATTCAATATTAATGAGACAGTACTAGACATGGGCC 1769  
QY 1621 aactaaatcagatcctcgtgctcagcccgctactactagttcgtlaatatlgagaag 1680  
|||||  
Db 1770 AACTAAATCAGATCTCGTGCGCTGACCCGCTATTACTAGTTCTGTATATATGAGAG 1829  
QY 1681 agatctagcttcaggactcaatgtgcccctccctcatctgtctacaagaatctgtagatca 1740  
|||||  
Db 1830 AGATCTAGCTTCAGGACTCATTTGGCCCTCTCTCATCTGCTACAAAGAAATCTGTAGATCA 1889  
QY 1741 aagaggaacacagaataatgtcagaagaagaatgtcatcctgtttctgtatltgata 1800  
|||||  
Db 1890 AAGAGGAACCACTAATATGTCAAGCAAGAGAAATGTATCTCTGTTTCTGTATTTGATGA 1949  
QY 1801 gaaccgaagctgttacctcagagaaatatacaagcgtttctcccaatccagcttgagat 1860  
|||||  
Db 1950 GAACCGAAGCTGTGACTCTCAGAGAAATATACAAGCGCTTCTCCCAATCCAGCTGAGT 2009  
QY 1861 gcagcttgagatccagagatltccaagctccaacatcatgcaagcgtcaatgtctatgt 1920  
|||||  
Db 2010 GCAGCTTGAGATCCAGAGTTCCAAAGCCTCCAAATCATGACAGCATYCAATGGCTATGT 2069  
QY 1921 ttctgatagtttgcagctgtcagttgttctgtcagtgagggtgcatactgtgtacattcaag 1980  
|||||  
Db 2070 TTTTGTAGATTTCAGCTGTGTCAGTTTGTTCATGAGAGTGGCATACTGTGATCTTAAG 2129  
QY 1981 catlgagacacagactgactccttctgtctctctctcctcgtatataccttcaaacaaa 2040  
|||||  
Db 2130 CATTGAGACACAGACTGACTTCTTCTGTCTCTCTGTGATATACCTTCAAAACACAA 2189  
QY 2041 aatgtgtctatgaagacacactcaacctatcccatctcagaagaactgtctctcatgtc 2100  
|||||  
Db 2190 AATGTCTATGAAGACACACTCAACCTATTCCTCAGAGAAACCTGTTCATGATC 2249

QY 2101 gatgaaaccccaagctcatgaltctgggtgtgcacaactcgaacttctggaacagag 2160  
|||||  
Db 2250 GATGGAACCCAGCTGTATGATTTCTGGGGTGGCCAACTCAGACTTTTGGAAACAGAG 2309  
QY 2161 catgaccgcttactgaagtttctagttgtgacaagaacactgtgtatatacagaaga 2220  
|||||  
Db 2310 CATTGACCGCTTACTGAAGTTCTAGTGTGACAAAGACACTGTGATTTATAGAGAGA 2369  
QY 2221 caattatgaagatattcagcatactgtcgtgagttaaaaaacaatgccaatgaaagaag 2280  
|||||  
Db 2370 CAGTTATGAAGATATTTCAACATACCTGTGAGTAAAAACAATGCCATTTGAMCCAAAG 2429  
QY 2281 ctctcccaagatccaagacacccctagcactagggcaaaagcaatttaatgccaacca 2338  
|||||  
Db 2430 CTCTCCAGATTTCAAGACACCTTAGCCTAGGCAAAAGCAATTTATGCGCACACA 2487

Search completed: November 18, 2001, 01:27:15  
Job time: 29577 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2001, 13:22:11 ; Search time 27.33 Seconds

(without alignments)  
3263.004 Million cell updates/sec

Title: US-09-689-430-2

Perfect score: 7869

Sequence: 1 MQELSTGCFELCLRFCSA.....VWHLALNMEVLGCEADLY 1471

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A\_Geneseq\_0601.\*  
1: /SIDSR/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SIDSR/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SIDSR/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SIDSR/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SIDSR/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SIDSR/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SIDSR/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SIDSR/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SIDSR/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SIDSR/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SIDSR/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SIDSR/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SIDSR/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SIDSR/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SIDSR/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SIDSR/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SIDSR/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SIDSR/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SIDSR/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SIDSR/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SIDSR/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SIDSR/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1          | 7857   | 99.8        | 1471   | 18    | AAW23414    |
| 2          | 7775   | 98.8        | 1457   | 20    | AAZ21675    |
| 3          | 7771   | 98.8        | 1457   | 19    | AAW46246    |
| 4          | 7771   | 98.8        | 1457   | 19    | AAW44372    |
| 5          | 7674   | 97.5        | 1438   | 21    | AAAB01262   |
| 6          | 7642.5 | 97.1        | 1440   | 12    | AAAR12971   |
| 7          | 7616   | 96.8        | 1661   | 18    | AAW18670    |
| 8          | 7615.5 | 96.8        | 1516   | 9     | AAAP80255   |
| 9          | 7585   | 96.4        | 1424   | 22    | AAW48842    |
| 10         | 7574   | 96.3        | 1425   | 9     | AAAP80267   |
| 11         | 7571   | 96.2        | 1424   | 9     | AAAP80268   |

|    |        |      |      |    |           |                     |
|----|--------|------|------|----|-----------|---------------------|
| 12 | 7567   | 96.2 | 1424 | 10 | AAAP91169 | Sequence of 740 Ar  |
| 13 | 7416.5 | 94.2 | 2342 | 18 | AAW11422  | Active Factor VIII  |
| 14 | 7416   | 94.2 | 2351 | 18 | AAW10591  | Factor VIII:C (Phe  |
| 15 | 7416   | 94.2 | 2351 | 18 | AAW13496  | Factor VIII:C (Arg  |
| 16 | 7416   | 94.2 | 2351 | 20 | AAZ21676  | Factor VIII proteol |
| 17 | 7415.5 | 94.2 | 2344 | 18 | AAW11432  | Active Factor VIII  |
| 18 | 7415.5 | 94.2 | 2344 | 18 | AAW11410  | Active Factor VIII  |
| 19 | 7415   | 94.2 | 2345 | 18 | AAW11415  | Active Factor VIII  |
| 20 | 7414.5 | 94.2 | 2346 | 18 | AAW11421  | Active Factor VIII  |
| 21 | 7414.5 | 94.2 | 2346 | 18 | AAW11431  | Active Factor VIII  |
| 22 | 7414.5 | 94.2 | 2346 | 18 | AAW11434  | Active Factor VIII  |
| 23 | 7414   | 94.2 | 2347 | 18 | AAW11411  | Active Factor VIII  |
| 24 | 7414   | 94.2 | 2351 | 18 | AAW10592  | Factor VIII:C (Tyr  |
| 25 | 7413.5 | 94.2 | 2348 | 18 | AAW11409  | Active Factor VIII  |
| 26 | 7413.5 | 94.2 | 2348 | 18 | AAW11414  | Active Factor VIII  |
| 27 | 7413   | 94.2 | 2349 | 18 | AAW11424  | Active Factor VIII  |
| 28 | 7413   | 94.2 | 2349 | 18 | AAW11428  | Active Factor VIII  |
| 29 | 7413   | 94.2 | 2349 | 18 | AAW11430  | Active Factor VIII  |
| 30 | 7413   | 94.2 | 2349 | 18 | AAW11420  | Active Factor VIII  |
| 31 | 7412.5 | 94.2 | 2350 | 18 | AAW11413  | Active Factor VIII  |
| 32 | 7412   | 94.2 | 2351 | 6  | AAAP50059 | Human factor VIII.  |
| 33 | 7412   | 94.2 | 2351 | 18 | AAW10590  | Factor VIII:C (Phe  |
| 34 | 7412   | 94.2 | 2351 | 18 | AAW11425  | Active Factor VIII  |
| 35 | 7412   | 94.2 | 2351 | 18 | AAW11427  | Active Factor VIII  |
| 36 | 7412   | 94.2 | 2351 | 18 | AAW11435  | Active Factor VIII  |
| 37 | 7412   | 94.2 | 2351 | 18 | AAW11408  | Active Factor VIII  |
| 38 | 7412   | 94.2 | 2351 | 18 | AAW11416  | Active Factor VIII  |
| 39 | 7412   | 94.2 | 2351 | 18 | AAW11419  | Active Factor VIII  |
| 40 | 7412   | 94.2 | 2351 | 18 | AAW11329  | Native Factor VIII  |
| 41 | 7412   | 94.2 | 2351 | 19 | AAW46245  | Human factor III P  |
| 42 | 7412   | 94.2 | 2351 | 19 | AAW44373  | Human Factor VIII.  |
| 43 | 7412   | 94.2 | 2351 | 21 | AAW52537  | Human full-length   |
| 44 | 7411.5 | 94.2 | 2352 | 18 | AAW11423  | Active Factor VIII  |
| 45 | 7411.5 | 94.2 | 2352 | 18 | AAW11426  | Active Factor VIII  |

#### ALIGNMENTS

|          |   |                             |
|----------|---|-----------------------------|
| RESULT   | 1   |                             |
| AAW23414 |   |                             |
| ID       | AAW23414  | standard; Protein: 1471 AA. |
| XX       |   |                             |
| AC       | AAW23414;   |                             |
| DT       | 08-APR-1998   | (first entry)               |
| XX       |   |                             |
| DE       | Human B-domain deleted factor VIII protein.                               |                             |
| XX       |   |                             |
| KW       | Post-translational regulatory element; PRE; enhancer II; intronless gene; |                             |
| KW       | surface antigen gene; cytoplasmic accumulation; targeted delivery;        |                             |
| KW       | near consensus splice sequence; blood coagulation factor; Factor VIII;    |                             |
| KW       | Factor IX.  |                             |
| XX       |   |                             |
| OS       | Homo sapiens.   |                             |
| XX       |   |                             |
| PN       | W09733994-A1.   |                             |
| XX       |   |                             |
| PD       | 18-SEP-1997.  |                             |
| XX       |   |                             |
| PF       | 10-MAR-1997;  | 97MO-US03561.               |
| XX       |   |                             |
| PR       | 11-MAR-1996;  | 96US-0683839.               |
| XX       |   |                             |
| PA       | (IMMU-) IMMUNE RESPONSE CORP.   |                             |
| XX       |   |                             |
| PI       | Bidlingmaier S, Ill CR;   |                             |
| XX       |   |                             |
| DR       | WPI: 1997-470874/43.  |                             |
| XX       |   |                             |
| DR       | N-PSDB; AAT73164.   |                             |
| XX       |   |                             |
| PT       | Vector for increased expression of intronless genes - comprises           |                             |
|          | intronless gene with at least one near consensus splice sequence, a       |                             |

promoter and at least one viral cis-acting post-transcriptional regulatory element

The presentencequence represents human B-domain deleted factor VIII. The cDNA encoding this protein also contains, 3' of the coding region, a post-translational regulatory element (PRE) of the Hepatitis B virus. PRE sequences have been shown to function *in cis* to increase the steady-state levels of surface gene transcripts by facilitating cytoplasmic accumulation of these transcripts. The above nucleic acid sequence is part of a novel vector, comprising an intronless gene containing 1 or more near consensus splice sequences operably linked to a promoter sequence so that the gene is transcribed in a cell. Intronless gene transcripts which contain near consensus splice site sequences are believed to get tied up in the nucleus of the cell where splicing occurs, rather than being transported to the cytoplasm where they can be translated into proteins. The PRE sequences are transcribed along with the gene, causing export of the gene transcript from the nucleus into the cytoplasm of the cell. The vector can be used to increase the expression of an intronless gene containing at least one near consensus splice sites, preferably cDNA encoding a blood coagulation factor, particularly Factor VIII or IX. The complex allows the targeted delivery of the vector to a specific cell, e.g. hepatocytes when the ligand is an asialoglycoprotein which binds the asialoglycoprotein receptor present on their surface.

Sequence 1471 AA;

|                            |        |               |           |              |
|----------------------------|--------|---------------|-----------|--------------|
| Query Match                | 99.8%; | Score 7857;   | DB 18;    | Length 1471; |
| Best Local Similarity      | 99.9%; | Pred. No. 0;  |           |              |
| Matches 1469; Conservative | 0;     | Mismatches 2; | Indels 0; | Gaps 0       |

|    |     |  |     |
|----|-----|--|-----|
| Qy | 1   | MOIELSTCEFFLCILRECCSARRRYLGAVELSNOMYOSDGLPVDARFPFRPRPKSPFN     | 60  |
| Db | 1   | myelstcefflcllrfcisatrryylgavelswmysdglpvdarfpfrprkspfn        | 60  |
| Qy | 61  | TSVYVKKTLEVEFYTHLNLNAKPRPMWGLLGPVIOAEVYDTVITLKNNAHPVSLHAV      | 120 |
| Db | 61  | tsvyvkkctlfvefthlnlknkprpwmglpvtioaevydtvitlknnahpvslhav       | 120 |
| Qy | 121 | GVSYWKASBEGAYEDDOTSQREKEDDKVPFGSHTYVMQVLKENGPMASDPCLITYSYLSH   | 180 |
| Db | 121 | gvsywkasegaeayeddqtsrqekeddkvpfgshcyvqvlkengpmasdpclitysy1sh   | 180 |
| Qy | 181 | VDLVYDNLNSGLIGALLNOREGSLAKEKQTLHKTLLFAYVDEKSKMHSKTSKLNQDRD     | 240 |
| Db | 181 | vdlvldnlsngllgallnveregslakekqtlhktllf1favdeeksvmhsetkslnmqdrd | 240 |
| Qy | 241 | AASARAPKPMHATVNGYVNRSLPGLGLCRRKSUVYMHVIGMGTPEVHSIFLEGHTFLVRNH  | 300 |
| Db | 241 | aasaraekpmhtvngyvnrslpglglcrrksuvymhvigmgtpevhs1fleghtflvrnh   | 300 |
| Qy | 301 | ROASLEISPTFFLVAQTLMLMDGQFLFCHTSSHOHDGMEAYYKVDSCPEEPOLRMKNNE    | 360 |
| Db | 301 | rgasleisptffltaqtlmlmdlgqfllfchssnhdgmeaykvvdscpeepqlrmkne     | 360 |
| Qy | 361 | EAEVDYDDLDTSEMHVYAFDDDNNSFQIOTSVAKKRPTVHHTIAAEEEDMYDAPLYLA     | 420 |
| Db | 361 | eaedydddltdsemvnyafdddnspsfqirtsvaknpkltvwh1aaeedwdyaprvla     | 420 |
| Qy | 421 | PDDRKYSCQYLNNGPQRIGRRYKKKRFMAVYDEFFKTRREALTOHSSGILGPLYLGVEVDTL | 480 |
| Db | 421 | pddrkyksqylngpqrigrkykkvrtmaydcelfkctreaqhesg1l1p1lygevgdtl    | 480 |
| Qy | 481 | LIIFKNOASRPYNIYPHGJTVDRPLYSHRLKGVKHLKDFPIELGELFFKMYTVEDGP      | 540 |
| Db | 481 | liifkngasrpynlphg1tdvprplysrllpkykhlkdf1lpgelfkykwrtvedgp      | 540 |
| Qy | 541 | TKSDPRCLTRYSSVNNNERDLASGLJPLLCIKESYVDQRGNOIMDKRNVILFSYFDE      | 600 |
| Db | 541 | tksdprcltryssvnnmerdlasglj1p1l1cykesyvdqrgn1mdkkrnv1l1fsyfd    | 600 |

|          |  |   |      |
|----------|--|---|------|
| Qy       | 601  | NRSWLTFTNIQRLPRLPNAGVQLDEPPEQASNMIMISNGVYFDSIQLSVCLHEFAVWYILS     | 660  |
| Db       | 601  | nrswwltftniqrllprrlpnagvqldeppeqasnmimisingvyfdsiqslvclhefawwyils | 660  |
| Qy       | 661  | IGACQDFLSVFSFGCTFFHKMWYEDTLTLPPFSGETVMSMENGLWILGCHNSDFNRG         | 720  |
| Db       | 661  | igacqdfslsvfsfgctffhkmyedttltpfsgetvmsmenglwllgchnsdfnrg          | 720  |
| Qy       | 721  | MTALIKVSSCDKNTGDYEDSYEDISAYLASKNNAIEPRFSQNSRHPSTROKOFNAPP         | 780  |
| Db       | 721  | mtalikkvsscdkntgdyedsyedisyaylasknaieprfsqnsrhpsttrkqfnaapp       | 780  |
| Qy       | 781  | VLKRRQRITRTTTLQSDQEEIDYDDTISVEKKREDPDYDDDENOSRSPQKRRRHFFIA        | 840  |
| Db       | 781  | vlkrrqrirttttlqsdqeeidyddtisvekkredpdydddenosrpsqkrrrhffia        | 840  |
| Qy       | 841  | AVERLMDMGSSPHVLNRRAQSSVPOKRVYQOEFDGSGFTOPLRGELNEHLGLGP            | 900  |
| Db       | 841  | averlwdmgssphvlnrraqsavpokrvyqoefdgsgftoplrgelnehlglgp            | 900  |
| Qy       | 901  | YIRAEVEDNIWTFERNQASRPYSFYSSLSIYEDDROGAEPRKNPVKNETKTYFWKVOH        | 960  |
| Db       | 901  | yiraevedniwtfernqasrpyfsyslsiyeedrogaepkrnpvknetktyfwkvoh         | 960  |
| Qy       | 961  | HMAPTKDEDDCAKAAVPSDVLKEDVHSGLLGPLLCVHTNTLNAHROVYVOEALFFT          | 1020 |
| Db       | 961  | hmaptkdeedcakaavpsdvlkedvhsghllgpllvchntnl nahrovyyoealffft       | 1020 |
| Qy       | 1021   | IFEDTKSWYFTFENMERNCRAPCNITQMEDPPEKENVRFHAIINGYIMDPLGLVMAODORIR    | 1080 |
| Db       | 1021   | ifedtkswyftfennerncrapcnitqmedppekenyrfhaingyimdtplglvmaodqrir    | 1080 |
| Qy       | 1081   | WYLLSMGSGNEINHSHSGHVFYTRKKKEKYKMALYNLPGVFETVEMLPSKAGIWRRECL       | 1140 |
| Db       | 1081   | wyllsmgsneinhshisghvfytrkkkeymalynlpgvfetvmlpskagiwrvectl         | 1140 |
| Qy       | 1141   | IGEHHLHASMSTLFLVYSNRKCTPLGMSAGHTRFDOITASGQYGQWAPRLARLHVSISINA     | 1200 |
| Db       | 1141   | igehlhmstlflvysnrkcptlgmasghtrfdqitcasgqygqwaprlarlhyssina        | 1200 |
| Qy       | 1201   | WSTKEPFSMIKVDLAPMIHIGIKTQGAOKFSSLIYQFTIMYSLDGKKMOTYRCNSTG         | 1260 |
| Db       | 1201   | wstkepfsmikvdlapmihigiktqgarokfssliysqftimysldgkkmotyrgnstrg      | 1260 |
| Qy       | 1261   | TLMPFGVNDSSGIKHNIENFPPIIARIYRLRPHYSTRSTIRMLMGCDLNSCMPRLGME        | 1320 |
| Db       | 1261   | tlmptfgvndssgikhnienfpptiariyrlrphystirstirmllmgcdlnscmprlgme     | 1320 |
| Qy       | 1321   | SKAISDAQITASSVFTTNFATWSPSKALHLCQGRSNAMRPQVNNPKPMLQVDFKTKKVT       | 1380 |
| Db       | 1321   | skaisdaqitassvfttmfatwspskalhhqgrsnamrpqvnnpkpmlqvdfktnkvt        | 1380 |
| Qy       | 1381   | GVTTQGVKSLTSMYVKEFLISSQDGHQWTLFQNGKRVKVGQGNODSFTEPVNSLDPL         | 1440 |
| Db       | 1381   | gvttqgvkslltsmyvkeflissqdghqwtlfqngkrvkvgqgnodsftepvnsldppl       | 1440 |
| Qy       | 1441   | LTRYLRIRHPQSWHQIALRMEVLCEANDLY 1471                               |      |
| Db       | 1441   | ltrylrirhpqswvhoialrmevlceandly 1471                              |      |
| RESULT 2 |  |   |      |
| ID       | AAAY21675  |   |      |
| XX       | AAAY21675  | standard; Protein: 1457 AA.                                       |      |
| XX       | AAAY21675:   |   |      |
| XX       | 18-AUG-1999  | (first entry)   |      |
| XX       | Beta-domain deleted Factor VIII protein.                                 |   |      |
| XX       | Factor VIII protein; gene modification; gene therapy; clinical disorder. |   |      |

KW splicing pattern: RNA processing; gene regulation; beta-domain; human.  
XX Homo sapiens.  
XX MO9929848-A1.  
PN 17-JUN-1999.  
PD 25-NOV-1998; 98MO-US25354.  
XX 16-JAN-1998; 98US-0071596.  
PR 05-DEC-1997; 97US-0067614.  
XX (IMMU-) IMMUNE RESPONSE CORP.  
PA  
XX  
PI Bidlingmaier S, Gonzales JEN, Ill CR, Yang CQ;  
XX WPI; 1999-385602/32.  
DR N-PSDB; AAX82258, AAX82259, AAX82260.  
XX  
PT Genes and vectors exhibiting increased expression and novel splicing  
PT patterns, useful for expression of, e.g. beta-domain deleted factor  
PT VIII  
XX  
XX  
PS Disclosure; Page 72-78; 123pp; English.  
XX  
XX The invention describes novel genes and vectors exhibiting increased  
CC expression and novel splicing patterns. It provides a gene encoding a  
CC Factor VIII protein, that comprises one or more consensus or near  
CC consensus splice sites which have been corrected to increase expression.  
CC The method, DNA sequences and expression vectors can be used to increase  
CC the expression of a gene, especially a Factor VIII gene. Genes containing  
CC modified 5' and/or 3' untranslated regions have optimized expression  
CC levels and tissue-specific expression. The methods are used for  
CC identification and correction of consensus splice sites, addition of  
CC introns, optimization of 5' and 3' untranslated regions and increase in  
CC cytoplasmic RNA accumulation. Hence the DNAs are useful in gene therapy  
CC to treat a clinical disorder, to study RNA processing and/or gene  
CC regulation. The present sequence represents a beta-domain deleted Factor  
CC VIII protein.  
CC  
XX  
S0 Sequence 1457 AA:  
  
Query Match 98.8%; Score 7775; DB 20; Length 1457;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 1456; Conservative 1; Mismatches 0; Indels 14; Gaps 1;  
  
QY 1 MQEISNCFEFLCLDFCFSSATRRYYLGAVELSDMYMOSDLGELPYDARFPRVRKSPFFN 60  
DB 1 meIeIstcfcIcIrrfcsatrryIlgaveIswdymgsdIgeIpydarIfrprvksIfrn 60  
  
QY 61 TSVVYKKTLEVEFTVHLFNIAKPRPPWMLGPTIQAEVYDVVITLKNMASHVSLHAV 120  
DB 61 tsvvykktlIfeftvhlfnIakprppwmllgptIqaevydvvtlIknmashvslhav 120  
  
QY 121 GVSVMKASEGAEVYDDQTSQRKEEDKVPFGSGSHYVQVLEKNGPMASDPLCLTYSLSH 180  
DB 121 gvsvmkasegaeYddqtsqrkeedkvIpgshhyvqvIlekgpmasdpIcltysylsh 180  
  
QY 181 VDVLKDLNSGLIGALVLCREGSLAKEKTQTALHKLTLFAVDECKSMHSEKNSLMORD 240  
DB 181 vdvlkdlnsglIgalvIcregslakektqtalhkIltIfavdegksmhseknslmqdrd 240  
  
QY 241 AASRAMPKMHTVNGVYNSRLPGLIGCHRSVYWHVIGMTPEVHSTIELGHTFLVNH 300  
DB 241 aasrampkmthtvngvynsrIpglIghchrsvYwhvIgmtpEvhstIeLghTflvnh 300  
  
QY 301 ROASLETSPITFLTAQTLMLDGLFLFCHTSSHOGHGMKAYVVDSCPEEPOLRMKNE 360  
DB 301 rgaSletspItfltaqtlmldglfIlfchTsshohgmkaYvvdscpeepolrmkne 360  
  
QY 361 EAEDYDDLTFDSEMDVVRFFDDNSPSFTQIRSVAKKHPKTVWVYIAAEEDMDVAPLVLA 420

DB 361 eaedydddlTfdsemdvvrffddnspstfIqirsvakkhpkTvwvYiaaeedmdvapyIvla 420  
QY 421 PDDRSYKSOYLNNNGFQIRGKRYKKRYFMAVYDFEFTKTBREAIQHEGSLIGPLLYGEVDTL 480  
DB 421 pddrsyKsoylNnngfQIRgkryKkryfMavYdfEftkTbreaiQhegslIgpLlygeVdtl 480  
QY 481 LIIFRNOASREYNYIPHCITDVRPLYSRRLEKGVAKHLKDFPLPGELEIKYKWTYVEGDP 540  
DB 481 lIIfRnoasreYnyIpHCITdvrPlYsrrLEkGVakHLkdfPlpgeleIKyKwtYvegdp 540  
QY 541 TKSDPRCLTRYSSPFWNNERDASGLIGPLLCYKESYVDGONGNOMSDKRWVILFESVDE 600  
DB 541 tksdprcltryssPfwnnErDASglIgpLlcykesyvdGONGnomSDkRWvIlfEsVde 600  
QY 601 NRSWLTENIQRFLEPNPAGVOLDEPEFOASINMHSINGVEDSLQSVCLHEVAYWYLLS 660  
DB 601 nrsWlTenIqrflePnpagVoldepefoasInmHsIngvedslQsvclhevaywylls 660  
QY 661 IGAQTDLFISVFSSGTYFEKHKWYEDTLTLFPFSGSTVPMSENREGLMILGCHNSDFRRNG 720  
DB 661 IgaqtdlfisvFssgtyfeKHKwyEdtLtlfPpfsGstvPMsENreGLmIlGchnsdfrnrg 720  
QY 721 MTALLKVSQCDKNTGDYEDSYEDISAVLLSKNNAIEPRFSQNSRHRSTROKOFNATPP 780  
DB 721 mTallkvsqcdkntgdYedysyEdisavllsknnaIeprfsqnsrhrstrokofnatpp 780  
QY 781 VLKRHOREITRTTLQSDQEEIDYDDTISVENKKEDFDIYDEDEMSPRSFQKTRRHYFA 840  
DB 781 vlkrhoreitrttlQsdqeeIdyDDtIsvenkkEdfdIydeDEmsprsfQktrrhyfa 840  
QY 841 AVERLMDVGSSSPHVLNRNROSGSVPOFKRVYVQEFNDGSEFTQPLYGLNELHGLLGP 900  
DB 841 averlmdvgssspHvlnrnrosgsvpoFkrvyvQefndgseftQpLyglneLhglLgp 900  
QY 901 YIRAEVEDINIVATFRNQRSRPYSFSSLSIYEEDROGAEPKRNKFNENETKTFYFWKVOH 960  
DB 901 yiraevedinivAtfrnqrsrPysslsIyeedrogaepkrnkfnenETkTfyfwkvoh 960  
QY 961 HMAPTKDEFDCKAMAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAGHQTVYQEFALFT 1020  
DB 961 hmaptkdefdckamayfsdvdleKdvHsglIgpLlvchTntlnpaghqtvYqefalft 1020  
QY 1021 IFDETKSWYFTENENRNRACNTQMEDPTFEKWRFRFALINGYIMDTPLGLVMAQDOQIR 1080  
DB 1021 ifdeTKswyftEnenrnracntQmedptFEkwrfrfAlIngYImdtPlglvmaqdOqir 1080  
QY 1081 WYLLSMGSENENIHSHFSGHYFTVRKKKEEYKMALYNLPGVEFTVEMLPKAGIWRVCL 1140  
DB 1081 wyllsmgseneniHshfsgHyftvrkkKEEYkmaLynlpgveftvEmlpkagiwrVcl 1140  
QY 1141 IGEHLHAGMSTFLVYSNKCOTPLGMASGHIRDFOITASGOYGWAPKLARIHVSCTINA 1200  
DB 1141 IgehlhagmstflvysnkcotPlgmASghirDfoitASgoYgWapklarihvsctina 1200  
QY 1201 MSTKEPFWIKIVDILAPMIIHIGITQGAOKFSSLYISQIFIMSLDOCKKQYTRGNSGTG 1260  
DB 1201 mStkePfwikivdIlapmIiHigItqgaokfsslyisqifImslDckkqytrGnsGtg 1260  
QY 1261 TLMWFFGNVDSGSIKHNFNPPIIARVIRLAPHTVYSIRSTRLMELMGDLNMSCMPLGME 1320  
DB 1261 tlmwffgnvdsGsiKhnfnpPiIarvIrLaphtvYsIrstRlmeLmgdlNmscmplgme 1320  
QY 1321 SKAISDAQITASSYFTNNFATWSPSKARLILQSGNSAMARPOVNNPKEMLOVDFOKTYMKVT 1380  
DB 1321 skaISdaqItassYftnnfAtwspSkarLilqsgnsaMarpoVnnpkemlovdfoKtymkvt 1380  
QY 1381 GVFTQGVKSLLTSMVYKFEFLSSSDGHWLTFPQNGKGVKFKFGQGNQDSFTPVVNSLDPPL 1440  
DB 1381 gvftqgvksllTsmvYkfeFlsssdGhwLtfPqngkGvkfkFGqgnQdsftPvvnslDppl 1440  
QY 1441 LTRYLRIPHOSWHAQIALRMEVLGCEAODLY 1471  
DB 1441 ltrylrIphoswHaqIalrmeVlgceAodly 1471

Db 1427 ltrylrlnpgswlnqialrmelvgceagdlly 1457

RESULT 3

AAW46246

AAW46246: standard; Protein: 1457 AA.

06-AUG-1998 (first entry)

Human factor VIII beta-domain deleted SON deletion protein sequence.

Replication defective; recombinant retrovirus; RRV; therapeutic protein; haemophilia; thrombosis; hypercoagulable disorder; liver disease; human; hepatitis; thalassemia; phenylketonuria; Lesch-Nyhan syndrome; diabetes; cystic fibrosis; Duchenne's Muscular Dystrophy; hypercholesterolemia; hypopituitarism; adenine deaminase deficiency; HIV infection; anaemia; Guacher's syndrome; high blood pressure; Alzheimer's disease; autoimmune; inflammatory disease; factor VIII.

Homo sapiens.

WC9800541-A2.

08-JAN-1998.

02-JUL-1997; 97WO-US11784.

04-JUN-1997; 97US-0869309.

03-JUL-1996; 96US-0645601.

13-AUG-1996; 96US-0696381.

(CHIR ) CHIRON CORP.

Allen JR, Barber JR, Boder M, Chang SMW, Chong K;

De La Vega D, Depoloni, Greengard J, Hsu DC, Ibanez CE;

Jolly DJ, Lee R, Mittelstaedt DM, Prussak CE, Respass JG;

NP1; 1998-086966/08.

N-PSDB; AAV19581.

New replication defective recombinant retro-viruses - which can be

administered to provide long term systemic expression of therapeutic

protein in blood, useful in, e.g. treating hyper-coagulable

disorders

Example 28; Pages 213-217; 272pp; English.

This is the beta-domain deleted SON deletion protein of human factor

VIII. The encoding DNA is used to construct recombinant retroviral

vectors expressing human factor VIII. The invention provides the

preparation of replication defective recombinant retrovirus (RVV)

expressing a therapeutic protein. The RVV preparation is resistant to

degradation by human complement and is capable of inducing long term

systemic expression of the therapeutic protein when administered

intravenously to a human. The long term systemic expression results in a

measurable level of the therapeutic protein being produced in the blood

of the human for a period of at least 30 days after the administration of

the RVV vector preparation. RVV's can be used for in vivo delivery of

Query Match 98.8%; Score 7771; DB 19; Length 1457;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 1456; Conservative 0; Mismatches 1; Indels 14; Gaps 1;

QY 1 MOLEISTCFCLLRFCFSATRRYILGAVELSWDMOSDLGELPVDARPPRRVRSFPFN 60  
DB 1 mqlelstcfcllrrfcfsatrryilgavelswdmgsdglpvdarfprrvksfpfn 60  
QY 61 TSVYKKTLEVEFTVHLFETIAKPRPMGLGPTLOAEVDYDTVTLLNMASHPSLAV 120  
DB 61 tsvykktlveftvhlfiakprpmmglgptloaevdydtvtllnmashpslhav 120  
QY 121 GVSVMKASGARYDDQTSREKEDKVPFGSGHTVMOLKENGPMASDPLCLTYLSH 180  
DB 121 gvsvmkasgareyddqtsrekedkvfgsghtvmolkengpmasdpclctysylsh 180  
QY 181 VDLVLDNLGILGALLVCEGSLAKEKTYTLAKFTLLRAVDEGKSMHSETRKNSLMQDR 240  
DB 181 vdlvldnlgilgallvcegslakektqtlhkfillfavidegkswshetknslmqdrd 240  
QY 241 AASARAMPKMTNVNGVYVNSLPLGLGCHRSYVMHVIQGTPEVHSIFLEGHTFLVFNH 300  
DB 241 aasarampkmhtnvngvyvnslpjlgchksyvmhviqgmtphevhsifleghctflvfnh 300  
QY 301 ROSLEISPTFLTAOTLLMDLGOELLFCHISSHODGMEAVYKVDSCPEEPOLRMKNE 360  
DB 301 rgasleisptfltaotllmdlmgfllfchissdqhdgmeavkvdscpeepqlrmkne 360  
QY 361 EAEEDDDDLTDEMDVYRRDDNSPFTOIRSVAKKHRTWVHTAAEEDMDYAPLYLA 420  
DB 361 eaeeddddltdemdvrrddnspsftoirsvakkhrtwvhtaeeedmdyaplyla 420  
QY 422 PDDRSKSOYLNNQPORIGRKTKKVFMAVTEPTEKTRALIOHESGILGPLYGVGDTL 480  
DB 422 pddrsksoylnnqporigrktrkvvfmayteptektralohehsgilgpllygvgdttl 480  
QY 481 LIIFKQASRPYNIAPHGTTDVRPLYSRRLPKGVNHLDFPLIPGEIRKMYTATVEDGP 540  
DB 481 liifkqasrpyniaphgttdvrplysrllpkgvnhldfplipgeirkywtatvedgp 540  
QY 541 TKSDFRCLTRYSSVNNERDLASGLIGPLICYESVDGQNMISKRNVILFSVDE 600  
DB 541 tksdfrcrltryssvnnerrdlasgligpllcykesvdqngnmskrnvilfsvde 600  
QY 601 NRSWITENIOFELPNPAGVOLEDEPEFOASNTMHSINGVFESLQSLVCLHEVAYWYLS 660  
DB 601 nrswiteniolfelpnpagvoledepefoasntmhsingvfeslqslvclhevaywyls 660  
QY 661 IGAQTDFLSVFSGYTFKKHWYEDTLTLPFSGEFTVMSMENPGLWILGCHNSDFRNRG 720  
DB 661 igatdflsvfsgytfkhwedyedtltpfsgftvmsmenpgllwlgchnsdfrnrg 720  
QY 721 MALLAKVSSCDKNTGDEYEDYEDISAYLLSKNNALIEPRSSONSRRHSTKQKOFNAPP 780  
DB 721 mallakvsscdkntgdeyedyedisayllsknnalierprssonsrrhstkoqofnapp 780  
QY 781 VKRHOIRETRFTTLOSDEIDYDPTISVEMKKEPDFIYDDENSPSPFOKTRRHYFA 840  
DB 781 vkrhoiretrfttlosdeidydpitvsemkkepdfiyddenspspfoqktrrhya 840  
QY 841 AVERLMDYGMSSSPVLRNRAOSGSPVQFKVVFQETFDGFTOPLRGLNELHGLGLP 900  
DB 841 averlmdygmssspvlnrnraosgspvqfkvvfqtfdgftoplrglnehlhglglp 900  
QY 901 YRAEEDINIMTFRNQASRPYSFYSLSIYEEDROGAEPRKRVKPNETKTYFKWQV 960  
DB 901 yraeedinimtfngasrpyfsfyslsyeeedrogaeprrkvkpnetyfkvwq 960  
QY 961 BHAPTKDERDCKAMAYFSDVDEKDVHSGILGPLLYVCTNTLNAHGNQVVOEFALEFT 1020  
DB 961 bhaptkderdckamayfsdvdekdvhsgilgpllyvctntlnahgnqvvoefaleft 1020  
QY 1020 947 hmapktdetdckawayfsdvdekdvhsgilgpllvctntlnahgnqvrvlvgfalf 1006



```
OY 1021 IFDETRKSWFTEENMERNCRAPCNIOMEDPTEKENYREHAINGYIMDTPLGLVMAODRIR 1080
DB 1007 lfdetksywfteemerncraponimmedptfkenyrfinaingyindt.pglvmaodqrir 1066
OY 1081 WYLSMGSNENIHSHSGHVFYRKKEEYKALYNLYPGVEYEMLPKSAIGIWRVECL 1140
DB 1067 wylismgsnenishshghvftvrkkeeykmalynlypyvfevemlpksagiwrvectl 1126
OY 1141 IGEHHAAMSTFLVYSKCKQTPICMASGHIRDPOITASSGOYGQAPKLARLHYSGSINA 1200
DB 1127 lgehhaamstflvyskckqtpicmasghirdqitassgyqgwapklatrhyssina 1186
OY 1201 WSTKEPFMIKVDLAPAIHIGIKQCAROKFSSLYISQFIIMYSLDCKKQTYGNSGTG 1260
DB 1187 wstkepfmikvdlapailhigikqcarokfsslyisqfiimysldckkqtygnsbtg 1246
OY 1261 TLMVFFGVNDSSGIRKHNIFNPPIARVIRLHPHYSIRSTLRMELMGCDLNSCMPLGME 1320
DB 1247 tlmvffgvndssgirkhnifnppliarvirrlhphysirstlrmeimgcdlnscmplgme 1306
OY 1321 SKAISDAITASSYFTNNFATWSPSKARLHIOGRNAMPQVNNKEMLOVDFOKTMYT 1380
DB 1307 skaisdaigtassyftnmfatwspskarlhgirsnawrpqvnnpkewlqvdfoktmvt 1366
OY 1381 GVTTQGVSLTSMYKREFLISSSODGHQWTLFPONGKVKYFQGNODEFTFPVNSLDPEL 1440
DB 1367 gvttqgvsltismykrfelfissodghqwtlfpngkvkvyfgnqdsfcpvnsldppl 1426
OY 1441 LTRYLRHPQSWHQIALRMEVLGCEAODLY 1471
DB 1427 ltrylrhpgswbqialrmevlgceaqdly 1457

RESULT 4
AAW4372
ID AAW4372 standard; Protein: 1457 AA.
XX
AC AAW4372:
XX
DT 20-JUL-1998 (first entry)
XX
DE Human Factor VIII SON deletion mutant.
XX
KW Factor VIII: blood clotting; haemophilia A; gene therapy;
KW retrovirus; vector; human.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN W09800542-A2.
XX
PD 08-JAN-1998.
XX
PE 02-JUL-1997; 97WO-US11785.
XX
PR 04-JUN-1997; 97US-0869309.
PR 03-JUL-1986; 96US-0645601.
PR 13-AUG-1996; 96US-0696381.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Allen JR, Barber JR, Boder M, Chang SM, Chong K;
PI De La Vega D, Depolo NJ, Greengard J, Hsu DC, Ibanez CE;
PI Jolly DJ, Mittlestaedt DM, Prussak CE, Respass JG;
XX
DR WPI; 1998-086967/08.
DR N-PSDB; AAV15338.
XX
PT New replication defective recombinant retroviruses - which express B
PT domain-deleted human factor VIII or human factor IX for the
PT treatment of haemophilia
XX
```

```
PS Claim 5; Page 175-180; 236pp; English.
XX
CC This polypeptide comprises the B domain deletion mutant SON of
CC human Factor VIII. The SON mutant is created by fusing Ser-743 to
CC Gln-1638 of native Factor VIII (see AAW4373) to form a Ser-Gln-Asn
CC (SON) link between the A2 and A3 Factor VIII domains. A DNA
CC sequence encoding the SON deletion mutant is provided in AAV15338.
CC When compared to plasmid-derived Factor VIII, the SON deletion does
CC not influence the in vivo pharmacokinetics, but the reduced size of
CC the molecule appears to decrease proteolytic degradation. The
CC invention relates to preparations of replication defective
CC recombinant retrovirus (RV) expressing a B domain-deleted human
CC Factor VIII protein, where the recombinant RV is capable of
CC infecting human cells, is resistant to degradation by human
CC complement and is capable of inducing long-term (at least 30 days
CC and up to 6 months or longer post-injection) systemic expression of
CC Factor VIII when administered to a haemophilia A patient.
XX
SQ Sequence 1457 AA:

Query Match 98.8%; Score 7771; DB 19; Length 1457;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1456; Conservative 0; Mismatches 1; Indels 14; Gaps 1;

OY 1 MOIEISTGCFICLIRFCFSATRRYVLGAVELSMWDMOSDLGELPYDARFPVRPSPFN 60
DB 1 mqlststfcilcrlfcsatrryylgavelswdmgdgelpydatfprvpksfpfn 60
OY 61 TSVYVKKTLEVEFTVHLEFNIAKPRPMMGLGPTVQAEVYQTVVYTLKNMASHPSLAV 120
DB 61 tsvyvkkklfevftchlniakprpmmgllgptqaevydtvtvltknmshpslshav 120
OY 121 GVSYWKASEGAEYDDQTSQREKEDDKVPGSGHYYWVOYLKBNBPMASDPLCLTYSTYSH 180
DB 121 gvsywkasegaeyddqtsqrekeddkvpgsgshcyvgyvylkbnpmasdpcltystysh 180
OY 181 VDLVKDNLNGLIGALLVCREGSLAKEKTOQLHKFLLFAVDEGKSMSEFNKSNLMOORD 240
DB 181 vdlvkdnlsnlgallvcregslakekctlhkfllfavidegksmwsefnksnsmord 240
OY 241 AASARAMPKMTGVNGYVNRSLPGLGCHRSYVWHVIGMTTPEVHSIFLGGHFFLVNHH 300
DB 241 aasarawpkmtlvgnyvrslpglgchrsyvwvhvigmtpvehsiflghftlvnrh 300
OY 301 RQASLEISPIITFLRAQTLMDLGGFLFCHTSSHOHDMEAIVKVDSCPEEPQLRMKNE 360
DB 301 rqa sleispiitflraqtlmdlgqflfchissbqhdmeayvkvdsceepqlrmkne 360
OY 361 EAEDYDDDLTDSMDVYRFDDDNSPSFTQIRSVAKKHKPTWVHTYIAAEEDMDVAPLVLA 420
DB 361 eaedydddltdsmdvyrfdDNSpsftqirsvakkhpktwvhtyiaaeedwdyaplvla 420
OY 421 PDDRSYSQSYLNNNGFORIGRKYKKVRFMAVYDEFEKTRKATQHSGLIGPLLYGEVGDTL 480
DB 421 pddrsysqsylnngforigrkykkvrfmayvdeftktrcatqhsagllgpllygevgdtl 480
OY 481 LITFRNQSAPRYNIYPHGITDVRLYSRRLPKYKHLKDPILPGEITFKYKWTYVEDGP 540
DB 481 litfrnqsapryniypghitdvrlpysrrlpkykhlkdpilpgeitfkykwtvvedgp 540
OY 541 TKSDPRCLTRYSSFVNNERDLASGLIGPLLICKESYDQDGNQIMSKRNVITLFSVDE 600
DB 541 tkdprcltryssfvnnerdlasgllgpllickesydqdgngimskrnviltfsvde 600
OY 541 tkdprcltryssfvnnerdlasgllgpllickesydqdgngimskrnviltfsvde 600
DB 541 tkdprcltryssfvnnerdlasgllgpllickesydqdgngimskrnviltfsvde 600
OY 601 NRSWYLTENIQRLPNPAGVLEDEPEFQASNIMHSINGYVDSQLSCLCHEVAWYIIS 660
DB 601 nrswylteniqrlpnpagvledepefasnimsingyvdsqslsclchevaywylis 660
OY 661 IGAOTDPLSVFSGYTRKHKWYEDTLTLFPFSGEYFVMSMENPGMLWLGCHNSDFRNNG 720
DB 661 igatdplsvfsgytrkhwedyedtltlfpfsgetyfvmsempgmlwlgchnsdfnrng 720
```

```

QY 721 MTALIKVSSCDKNQDGYEDSYEDISAYLLSKNNALIEPSSFSONSRHPSTROKQFNATPP 780
    |||||
    |||||
    |||||
Db 721 mtaalikvsscdknqdyedyedisayllsknnalepssfsqn-----pp 766
QY 781 VLKRRQRETRRTTLOSDOEIDYDDTISYEMKKEDFDIYDEDNOSPRSPQKTRRHYYTA 840
    |||||
    |||||
    |||||
Db 767 vlkrqretrrttlosdoeidyddtisvemkkedfdiydednqsprstqktrrhyyta 826
QY 841 AVERLMDYGMSSSPHYLRRRAQSGSVQPKVVFQETGSGFTQPLYRGELNHELGLGP 900
    |||||
    |||||
    |||||
Db 827 averlmdygmsssphylnrraagsvqpfkvvfgetfgsfqplryrgelnhehlglp 886
QY 901 YIRAEVDNINWYFRNQAQSPYSFYSLSLYEDDQGAEPKRNKVPNETKTYEMKVOH 960
    |||||
    |||||
    |||||
Db 887 yiraevedninwfrnqaqspysfyslslyeedqrgaepkrnfvkpnetkyfemkvoh 946
QY 961 HMAPRKDFDCKAMAYFSVDLEKDVHSGLLIGPLVCHNTNLNPAAGROVTVQEFALFT 1020
    |||||
    |||||
    |||||
Db 947 hmaprkdfdckamayfsvdlekdvhsglligplvchntlnpahrqvtvqefalf 1006
QY 1021 IEDETKSWYFTENMERNCRAPCNIQMEDPTFEKENYRFHAINCYIMDLPLGLVMAQDQRT 1080
    |||||
    |||||
    |||||
Db 1007 ifdetkswyftemerncrapcnlqmedptfekenrfhaincyimdlplglvmaqdqr 1066
QY 1081 WYLLSMGSNENIHSHFSGHVFTRKKEEYKMAALNLYGVETVEMLPDSKAGIMVECL 1140
    |||||
    |||||
    |||||
Db 1067 wyllsmgsnenihshfsghvftrkkeeykmaalnlvgvetvemlpdkagimvecl 1126
QY 1141 IGENHAGMSTFLVYSNCKQPTPLGASGHIRDFOITASGOVGMPPKTLARLHYSGINA 1200
    |||||
    |||||
    |||||
Db 1127 igenhagmstflvysnckqptplgassgihirdfoitasgovgmppktlarlhyssina 1186
QY 1201 WSTKEPFSMIKVDLAPMIHIGIKTOGARQKFSLSYISQFIIMYSLDGKRWQTYRGNSGT 1260
    |||||
    |||||
    |||||
Db 1187 wstkepfsmikvdlapmihigiktogarqkfslyisqfiimysldgkwwqtyrgnsgt 1246
QY 1261 TLMVFGVNDSSGIRKHNIFNPPIIARYIRLHPTHTYSIRSTLRMELMGCDLNSGSMPLGME 1320
    |||||
    |||||
    |||||
Db 1247 tlmvfgvndssgirkhnifnppliiaryirlhpthtysirslrmelmgcdlnscsmplgme 1306
QY 1321 SKAIDAOITASSYFTNMFPATSPKARLHLQGRSNAMRPVNNPKEMLOVDFQKTMKYT 1380
    |||||
    |||||
    |||||
Db 1307 skaidaoitassyftnmfpatspkarhlhlogrsnamrpvnnpkemlovdqkcmkyt 1366
QY 1381 GVTTCGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNODSFTPVVNSLDPP 1440
    |||||
    |||||
    |||||
Db 1367 gvttcgvkslltsmyvkeflissqdgqhwclffngkvkvfgnqdsftpvvnsldpp 1426
QY 1441 LTRYLRHPQSWHQAIALRMEVLGCEADPLY 1471
    |||||
    |||||
    |||||
Db 1427 ltrylrhpqswvhqialrmevlgceadply 1457

```

```

PF 08-DEC-1999; 99WO-US29169.
XX
XX 10-DEC-1998; 98US-0209916.
XX
PA (FAR ) BAYER CORP.
PI Cho M, Chan SY, Kelsey W, Yee H;
XX WPI: 2000-431311/37.
XX
PT Producing cells expressing a protein having factor VIII procoagulant
PT activity especially, human factor VIII in an industrial scale, involves
PT expressing a vector comprising a sequence coding for factor VIII in
PT human cells
PS Claim 7, Fig 1, 27pp: English.
XX
XX Producing cells expressing a protein having factor VIII procoagulant
XX activity, comprises contacting the cells with a vector comprising a
XX selectable marker and a sequence coding for the protein having factor
XX VIII procoagulant activity operably linked to a promoter. The cells
XX are then selected and individual clones expressing high levels of the
XX protein are isolated from the selected cells. The cells produced by
XX the method are not only useful for producing protein having factor
XX VIII procoagulant activity but also for producing adenovirus and
XX adeno-associated virus strains for gene therapy. The advantage of
XX having cells producing protein with factor VIII procoagulant activity
XX is that factor VIII protein can be produced on an industrial scale
XX in the range of 2-4 pg/cell/day. Human Burkitt's lymphoma (HKB)
XX cells provide a protein-free production system to produce not only
XX B-domain deleted factor VIII but also other therapeutic proteins. The
XX vector used in the method preferably comprises B-domain deleted
XX factor VIII (BDD-FVIII), a transcriptional unit for BDD-FVIII and a
XX selectable marker, dihydrofolate reductase (dhfr). In addition, a
XX terminal repeat sequence from Epstein-Barr virus is inserted into the
XX vector to increase integration efficiency.
SQ Sequence 1438 AA:
Query Match 97.5%; Score 7674; DB 21; Length 1438;
Best local Similarity 99.0%; Pred. No. 0;
Matches 1438; Conservative 0; Mismatches 0; Indels 14; Gaps 1;
QY 20 ATRRYLAGAVELSWDMQSDLGELPYDARPPRPVKSFPFNTSVYKTLFVEFTVHLEN 79
    |||||
    |||||
    |||||
Db 1 atrrylagavelswdmqsdlgelpydarpprvpksfntsvyktlffveftvhlfn 60
QY 80 IAKRPKPMGLIGPTIQAEVYDTVVITLKNMASHYSLHNAVSVWKSSEGAHEYDDQISQ 139
    |||||
    |||||
    |||||
Db 61 iakrppkmglligptlqaevydtvvitlknashpvslhavsvwksasegaeyddqtsq 120
QY 140 REKEDKVPFGSGHYVWQVVLKENGPMASDPICLTYSYLSHVDLYKDNLSGLIGALLVCR 199
    |||||
    |||||
    |||||
Db 121 rekedkvfpgshyvwvylkengpmasdpiclysyshvdlykdlnsgllgallvcr 180
QY 200 EGSIAKEKTQTLLHKFTLLPAVDEGKSWHSETKNSILMODRAASAARAMPKMTYGVYNR 259
    |||||
    |||||
    |||||
Db 181 egslakektqtlhkftllfavdeqkswhsetknsimgdrdaasarawpmkhtvngynr 240
QY 260 SLPLGLGCHRRKSYWVWVIGMGTPPEVHSTFLGSHFTLVANNHQALESPTFLTAQTL 319
    |||||
    |||||
    |||||
Db 241 slpllgchrrksywvwigmttpevhsflgshftlvnnhrgaselsptfltaqtl 300
QY 320 MDLQGLFCHSHSSHODMEAYVVKVDSCEPSPOLRMKNNEEAEDYDDDLTDSMDVYRF 379
    |||||
    |||||
    |||||
Db 301 mdlqgflfchshshodmeayvkvdsceepqlrmkneeaedyddldtmdvyrf 360
QY 380 DDDNSPFTQIRSVAKRHKPTWVHYTAAEEDMDVAPLVADDBRSYKSYQLNNGPQRIG 439
    |||||
    |||||
    |||||
Db 361 dddnspftqirsvakrhpktwvhytaeedwdvaplvlapddrsyksqylngpqr 420
QY 440 RRYKKVRPMAYIDETKTRDAIQHESGILGPIILLYGVBGTLLIIRKNOASRYNIYRPGI 499

```

```

Db 421 rkykvrifmayldetfkrcaiqhesgllgpllygvegdcllllffkqaspryllylpgl 480
QY 500 TDVRLPLSRRLPKGVKHLKDPILLPELFEKTKWIVTVADGPTKSDPRLTRYSSFWME 559
Db 481 tdlvrlplsrilpkvgkhkldfllpgelfkykvtvedgdplcksdprcltryssfwme 540
QY 560 RDLASGLLGPLICYSKESVDGRGNIMSDKRNVLFSVFEDNRSWYLTENIORLPNPAG 619
Db 541 rdlasgllgpllicykesvdgrgnimskdnvllfsyfdentrsywlylenqrlilnpag 600
QY 620 VQLEDPPEFOASINMHSINGVYFDSLOLSCLEHVAWYILSIGAQOPLSYFSGYTFKH 679
Db 601 vqledppegasimhsingyvfdsloqlsvclhevawylisigqtdlsvfsgytfkh 660
QY 680 KMYVEDTLTFPESGEYFMENMPGLMILCCHNSDPNRCMTALLKXSCDKNTGDYE 739
Db 661 kmvyedtltlfrsgetvfmempglwllqchnsdfrnrgmtallkxscdkntgdye 720
QY 740 DSYEDISAYILSKNNATEPRFSNSRHPSTRQKQFNATPPVLKRHOEIRRTLQSDOE 799
Db 721 dsyedisayilsknnaleprsfsgn-----ppvlkrhqeirtlclgsqde 766
QY 800 EIDVDITISVEMKKEDEPDYDEDENSPSPQKTRHYFIAVERLMDYGMSSPHVLRN 859
Db 767 eidvdtisvemkkedfdyde dengsprsfqktrhyfiaaverlmdygmssphvlrn 826
QY 860 RAQSGVPQFKKVVQFETDGSFTQPLRGELNHLGLGFIYRAVEDNIMVTFRNQAS 919
Db 827 raqsgvvpfkvvqfgetdgsftcpqlyrgelnhlgllyiraavednimvtfirngas 886
QY 920 RPYEYSLSLYEEDROGAPRKNFVKNPTKTYFMKVQHMMPTDEPFCKMAAFSD 979
Db 887 rpyeyslslyeedrgaaprknfvkpnctkyfwkvqghmmpctdeidckawaytsd 946
QY 980 VDLEKDVHSGLIGPLVCHTNTLNPAGRQVTVQEFALFTIPDETSMWTFEMNERCR 1039
Db 947 vdlekdvhsglilgplvchtntlnpahrqvtvqefalfflfdetkswytemenncr 1006
QY 1040 APCNIQMEDPFEKNRYPHAINGYIMDTLPGLVMAODRIRWYLLMGSNENHSHFSG 1099
Db 1007 apcnlqmedpfekenyrfhaiingyimdltplglvmaodqrlrwyllmgsnenhshfsg 1066
QY 1100 HVFVRKKEEKMLNLYPCVEFTEVEMLPKAGIWRRECIIGHILHAGNSTLFLVYSNK 1159
Db 1067 hvfvrkkeeekmallylpcvfevtemlpkaglwrvcecllghlhaqmstlflvysnk 1126
QY 1160 CQTPLGMAHGHIRDFQITASGOYQMAPKLARLHYSGSINAMSTKEPFWIKVDLAPMI 1219
Db 1127 cqtprlgmasghlirdfqitlasgqyqwapklarlhysgsinawstkepfswikvdlapmi 1186
QY 1220 IHGIRKTGAKROKSSLIYSOFITMISLDGKKWQTYRGNSGTGLWFGNVDSSGIRAHNF 1279
Db 1187 ihgirktagarkssliysqfilmysldgkkwqtyrgnstgclmwffgnvdsagikhnlff 1246
QY 1280 NPPIIARIIRLHPHYSIRSLRMEILMGCDLNSCSMPRGMSKSAISAQITASSYTFNMF 1339
Db 1247 nppliarlirhphysirslrmeilmgcdlnscsmprgmsskalsdaqitassytfnmf 1306
QY 1340 ATWSPSKARLHLQGRSNAMRPQVNNPKEMLOVDFOKTKWYGVTTQCGKSLTFSMYKEF 1399
Db 1307 atwspskarlhlqgrsnamrpqvnnpkewlqvdfqkcmkvsgvtgyskalltemykef 1366
QY 1400 LISSQDGHOWTLFFONGKAVFOGNOGSTPPVNSLDPLLTRYLRHPOSWHQAIALR 1459
Db 1367 liessqdghwtlffngkvkfsgngdsftgvnsldpplltrylrhpsqwhqialr 1426
QY 1460 MEVLGCEAOADLY 1471
Db 1427 mevlgceadly 1438

```

RESULT 6

```

AAR1297L
ID AAR12971 standard; protein; 1440 AA.
XX
AC AAR12971;
XX
DT 02-OCT-1991 (first entry)
XX
DE Factor VIII:SQ.
XX
KW Factor VIII; B domain; haemophilia.
XX
PN MO9109122-A.
XX
PD 27-JUN-1991.
XX
PF 06-DEC-1990; 90WO-SE00809.
XX
PR 15-DEC-1989; 89SE-0004239.
XX
PA (KABI ) KABIVITRUM AB.
XX
PI Almstedt AB, Hellstrom EM, Larsson K, Lind P, Sandberg HI, Spira J;
PI Sydow-Backman M;
XX
DR WPI; 1991-208148/28.
XX
PT Recombinant human factor VIII deriv. deoxyribonucleic acid -
PT encoding protein comprising two chains linked by segment of B domain.
XX
PS Disclosure: Fig 1; 35pp; English.
XX
CC The protein is a fusion between Phe 742 and Ser 1637 of the factor
CC VIII protein (factor VIII:SQ). In order to produce a
CC factor VIII deletion derivative that can be produced in vivo and/or
CC in vitro, to a two chain protein consisting of polypeptide chains of
CC 90 kd and 80 kd, the amino acid sequences surrounding Arg 740 and
CC Arg 1648 have to be conserved in order to preserve the structural
CC requirements for correct cleavage. In this example, amino acids 743
CC to 1636 of the full-length factor VIII polypeptide are deleted. A
CC new polypeptide chain is obtd. where there are 14 amino acids
CC linking Arg 740 and Arg 1648. Of these 14 amino acids, the sequence
CC of the five N-terminal ones directly corresponds to the five amino
CC acids following Arg 740 in full-length factor VIII. Also, the sequence
CC of the 12 C-terminal amino acids of the above 14 amino acids fragment
CC directly corresponds to the 12 amino acids preceding Glu 1689 in full-
CC length factor VIII, thus creating a 3 amino acid overlap between the
CC N- and C-terminal regions of the B-domain.
CC The factor VIII deriv. is useful for treating haemophilia or
CC haemophilia A. It has the biological characteristics of plasma derived
CC factor VIII.
CC In order to index this example, the factor VIII:QD amino acid
CC sequence was retrieved from W08800831 (AAB80265).
CC The amino acid numbering in the above comments is reproduced from the
CC fig. description in the specification. Note that Arg 740 is Arg 742
CC in the indexed sequence, etc. Also, Asn 745 (N-terminal link overlap)
CC is Asp 747 in AAB80265, but indexed as Asn to reproduce the fusion
CC fragment as shown in fig.1, and Asn 745 (C-terminal link overlap).
XX
SQ Sequence 1440 AA:

```

Query Match 97.1%; Score 7642.5; DB 12; Length 1440;  
Best Local Similarity 98.7%; Pred. No. 0;  
Matches 1434; Conservative 1; Mismatches 3; Indels 15; Gaps 2;

```

QY 20 ATRRYVLGAVELSDVMQSDLGELPVDARPPRPKSPFMTSVYKKTLFVEFTVHLFN 79
Db 2 atrryylgavelswdvmsdldgelpvdarpprpkspfmtsvykkltlveftvhlfn 61
QY 80 IAKRPPEWMLGLPTIOAEVYDVTVITLKNASHPVSLHAAGVSYWKASEGAEYDDQTSQ 139
Db 62 iakrppewmlglptioaevydvvtvitlknashpvsilhavgsywkasegaeyddqtsq 121

```

QY 140 REKEDKVPGGSHYVWQVLKENGPMASDPLCLTYSYLSHVDLVKDLNSGLIGALLVCR 199  
 |||||  
 Db 122 rekeddkvfpgshlywvwlkengpmasdpclctysylshvdlvkdlnsgllgallvcr 181  
 |||||  
 QY 200 EGSJAKEKTOHTHKFILLFAVEDEGSMHSEFTKNSLMQORDAASAAWAKMHTVNCYVR 259  
 |||||  
 Db 182 egslakekethlthkfillfavidegskwshetknsimgdrdaasarawkmhvngvvr 241  
 |||||  
 QY 260 S-LPGLICHRKSVYMHVIGMCTPEVHSIFLEGHTFLVRNHRQASLESPITFLAQTL 318  
 |||||  
 Db 242 slpglljchtrksywhvlgmgtpevhsifleghtflvrnhqaslelspitfltaql 301  
 |||||  
 QY 319 LMDJGFLFCHISSHQHDGMEAYKVDSCPEEPOLRMKNNEBAEDYDDLTDSEMDVVR 378  
 |||||  
 Db 302 lmdjgfillfchissqhgmeyaykvdscpeepqlrmkneeaedydddltdsemdlvvr 361  
 |||||  
 QY 379 FDDNSPFIOTRSYAKKHPRKTWVHYIAEEDMDVAPVLAPODDSYKSYLNNGPQRT 438  
 |||||  
 Db 362 fddnspsfllqtrsvakknprktwvhyiaeedwdyaplvlapddtsyksylnngpqrl 421  
 |||||  
 QY 439 GRKYKVFPMAYTDEFTFREALIOHESGILGPLYGEVDTLITFRKNOASRPYNIYPHG 498  
 |||||  
 Db 422 grkykvrfmeytdetfktrealiohesgllgpllygevdtllitfrkngasrpynlpyhg 481  
 |||||  
 QY 499 ITDVRPLYSRRLPKGVKHLKDFPILPGEIEFKYKWTYVEDGPTKSDPCLTRYSSFVNM 558  
 |||||  
 Db 482 ltvdrplysrllprkgvkhlkdfpilpgeiefkykwtvvedgptksdprcltryssfvnm 541  
 |||||  
 QY 559 ERDLASSGILGPLICRKESVQDQNGOIMSDKRNVLFSYFEDENRSYVLENTQRIPLPNA 618  
 |||||  
 Db 542 erdlasgllgpllicrkesvqdgngqmsdkrnvlfsyfdenrswyleenqrlfplpna 601  
 |||||  
 QY 619 GVOLEDPFOASNMHNSINGYFDSLSVCLHEVAYWYILSIGAOTDLSYFSSGYTRK 678  
 |||||  
 Db 602 gvqledpfgasnlnhsingyfdslqsvclhevaywylisigqtdlslsyfssgytrk 661  
 |||||  
 QY 679 HKMYEDTLTPFSGEYVPMSENPGLWILCNSDFRNKMTALLKVSCKDKNTGDYV 738  
 |||||  
 Db 662 hkmyedtltpfsgetvfmsempglwlilgchnsdfnrgmtallkvsckdkntgdyv 721  
 |||||  
 QY 739 EDSYEDISAVYLSKNNAITPRSESONSRHPRROKOFNTPRVLRKHOREIRRTQSQO 798  
 |||||  
 Db 722 edsyedisaayllsknnaileprsfsgn-----prvlrthrgelctcltqsdq 767  
 |||||  
 QY 799 BEIYDDTISYEMKKEDFDIYDEENQSPRSFOKTRHYFIAVERLMDYGMSSSPHVL 858  
 |||||  
 Db 768 eeiyyddtisvemkkedfdiydenngsprsfqktrhyfiaverlwdygmsssphvlr 827  
 |||||  
 QY 859 NRAQSGVPOKRVYFOEFTDGSFTQPLYRGLNHLGLGTYIRAEVEDNIMVTRNOA 918  
 |||||  
 Db 828 nraqsgvvpqfkkvfygeftdgsftqplyrgehlglglyiraavednimgvtrnqa 887  
 |||||  
 QY 919 SRPSFYSLSIYEBDOGAAPRKPNFKPNETKTYFWKVOHNMAPTQKPPCKAMAYFS 978  
 |||||  
 Db 888 srpsfysllsiyedqggaepkrkpnfknetktyfwkvoohnmaptkppckamayfs 947  
 |||||  
 QY 979 DVDLEKDVHSGILGPLICHTNTLNPARGROYVOEFALFEIIPDETSMYFTEMMERN 1038  
 |||||  
 Db 948 dvdlekdvhsgllgpllchntlnpargroyvofefalfefipdetksmyfitemernc 1007  
 |||||  
 QY 1039 RAPCNIQMEDPTFKENYFHAINGITMDTLPGLVMAODORIMYWLISGMSNENHSIHS 1098  
 |||||  
 Db 1008 rapcnlqmedptfkennyfhaingyimdtpglvmaodqrirwylismgsnenhsihes 1067  
 |||||  
 QY 1099 GHVFTVRKKEVKMALVNLVPGVFEVEMLPKAGIMWRKRECLIGELHNGMSTFLVYNS 1158  
 |||||  
 Db 1068 ghvftvrkkeekmalvnlvpgvfevemlpkagimwrrecligenlhaqgmstflvyysn 1127  
 |||||  
 QY 1159 KCQTPPLGNASGHIRDFQTASAGYQOMAPKLARLHYSGSINAMSTRKEPSWIKVDLAAM 1218  
 |||||  
 Db 1128 kcpqlgnasghirdfqtaagsgyqwapklarlhysgsinamstkepswlvkdlaam 1187  
 |||||  
 QY 1219 ITHGKTOGAKORFSSLYISQFIWYSLDGKKWOTYRGNSTGTLMVFGNVDSGIRKNI 1278  
 |||||

Db 1188 |||||  
 |||hgktgarqkfsslylsqfllmysldgkkwqfyrngstclmwfignvdsqikhni 1247  
 QY 1279 FNPPIARYIRLRHPHYSTRSTLRLMELMGCDLNSGMPLGMSKAIISAQITRASSYFNM 1338  
 |||||  
 Db 1248 fnpplaryirlnrphystrstclrlmelmgcdlnsgmpigmeskaisdaqitlaassyfntm 1307  
 |||||  
 QY 1339 FATWSPSKARLHLQGRSNAMRPOVNNPKEMLOVDQKTMKVYGVTTQGVKSILTSMYKE 1398  
 |||||  
 Db 1308 fatwspskarllhlgfnsnwrpvnnpkewlqvdyqtkmkygvttqgvksiltsmyke 1367  
 |||||  
 QY 1399 FLISSQDGHQWTLFQNGKRVYFQGNODEFMPVNSLDPPLTRYLRNIHQSWYHQIAL 1458  
 |||||  
 Db 1368 flissqdgghqwtlffqngkrvkvfgngdsftrvnsldppltrylrnhqswyhgial 1427  
 |||||  
 QY 1459 RMEVLGCEADOLY 1471  
 |||||  
 Db 1428 rmevlgceadly 1440  
 |||||  
 RESULT 7  
 AAM18670  
 ID AAM18670 standard; Protein; 1661 AA.  
 XX  
 AC AAM18670;  
 XX  
 DT 10-AUG-1997 (first entry)  
 XX  
 DE Factor VIII -db695-HCII.  
 XX  
 KW Factor VIII -db695-HCII; heparin cofactor II; blood coagulation;  
 KW blood clotting; procoagulant; anticoagulant; antithrombotic;  
 KW haemophilia; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 731..760  
 FT /label= HCII  
 FT /note= "heparin cofactor II acidic domain"  
 PN WO9118315-A1.  
 XX  
 PD 22-MAY-1997.  
 XX  
 PE 13-NOV-1996; 96MO-EP04977.  
 XX  
 PR 13-NOV-1995; 95US-0558107.  
 XX  
 PA (IMMO ) IMMUNO AG.  
 XX  
 PI Voorberg JJ;  
 XX  
 DR WPI. 1997-289291/26.  
 DR N-PSDB; AAT69811.  
 XX  
 PT Hybrid Factor VIII with modified activity, comprises region from  
 PT donor anticoagulant or antithrombotic protein - useful for treatment  
 PT of coagulation disorders  
 XX  
 PS Claim 11; Page 52-60; 96pp; English.  
 XX  
 CC Factor VIII -db695-HCII (AAM18670) is a hybrid protein in which amino  
 CC acids 712-736 of Factor -db695 (Factor VIII del868-1562) B-domain  
 CC are replaced by amino acids 51-80 from the acidic region (and  
 CC potential thrombin-binding site) of human heparin cofactor II  
 CC (HCII). It is the expression product of Factor VIII -db695-HCII  
 CC DNA (AAT69811) in plasmid pCLB -db695-HCII. The hybrid protein, which  
 CC can be expressed using gene therapy techniques, has increased  
 CC procoagulant activity owing to the HCII acidic region, and can be  
 CC used to treat blood coagulation disorders such as haemophilia A.  
 XX  
 SQ Sequence 1661 AA;

Query Match 96.8%: Score 7616; DB 18; Length 1661;  
 Best Local Similarity 87.0%: Pred. No. 0;  
 Matches 1450; Conservative 5; Mismatches 10; Indels 202; Gaps 4;

QY 1 MQEILSTCFELCLLRFCSATRRYYLGAVELSWDMQSDGLGELPVDARFPPRVKSEPFN 60  
 1 meielstcfelcllrfcsatrryylgavelswdymqsdglgelpvdarfpprvksfpfn 60

QY 61 TSVYKKTLPEFTVHLFNIAKPRPPMGLGPTIOAEVDTVTITLKNMASHVSLHAV 120  
 61 tsvykktlpeftvhlfniaakprppmglgptioaevdtvtitlknmashvslhav 120

QY 121 GVSYWKASEGAEDDOTSOREKEDKVPGGSHYVQVLEKNGPMASDPLCLTYSLSH 180  
 121 gvsywkasegaeddotsoerekedkvpggshyqvleknngpmasdpclcltysylsh 180

QY 121 GVSYWKASEGAEDDOTSOREKEDKVPGGSHYVQVLEKNGPMASDPLCLTYSLSH 180  
 121 gvsywkasegaeddotsoerekedkvpggshyqvleknngpmasdpclcltysylsh 180

QY 181 VDLVKDLNSGLIGALVCREGSLAKEKTOPLHKEFLLFAVEDEGKSMHSETKNSLMDRD 240  
 181 vdlvkdlnsgligalvcregslakektotplhkefllfavedegksmhsetknsldmrd 240

QY 181 VDLVKDLNSGLIGALVCREGSLAKEKTOPLHKEFLLFAVEDEGKSMHSETKNSLMDRD 240  
 181 vdlvkdlnsgligalvcregslakektotplhkefllfavedegksmhsetknsldmrd 240

QY 241 AASARAPKMHYNGVYNSRSLPGLIGCHRSVYWHVIGMCTTPEVHSIFLEGHTFLVRNH 300  
 241 aasaraapkmyhngvynsrslpgligchrsvywhvigmttpevhsifleghtflvrnh 300

QY 241 AASARAPKMHYNGVYNSRSLPGLIGCHRSVYWHVIGMCTTPEVHSIFLEGHTFLVRNH 300  
 241 aasaraapkmyhngvynsrslpgligchrsvywhvigmttpevhsifleghtflvrnh 300

QY 301 RQASLETSPITTEFLAOTLLMDLGOFLFCHSSHQHDGMEAYVAVDSCPEEPOLRMKNE 360  
 301 rqauletspitteflaotllmdlgoflfchsshqhdgmeayvavdscpeepolrmkne 360

QY 301 RQASLETSPITTEFLAOTLLMDLGOFLFCHSSHQHDGMEAYVAVDSCPEEPOLRMKNE 360  
 301 rqauletspitteflaotllmdlgoflfchsshqhdgmeayvavdscpeepolrmkne 360

QY 361 EAEDYDDDLTDSEMDVYRFDDDNPSFQIIRSVAKKPKXTVWHYIAAEEDMDAPVLA 420  
 361 eaedydddltdsemdvyrfdndnpsfiqirsvakkpkxtvwhyiaaeedmdapvla 420

QY 361 EAEDYDDDLTDSEMDVYRFDDDNPSFQIIRSVAKKPKXTVWHYIAAEEDMDAPVLA 420  
 361 eaedydddltdsemdvyrfdndnpsfiqirsvakkpkxtvwhyiaaeedmdapvla 420

QY 421 PDBRSVSYQLNNGPQIRGRKKRVPMAYTDETFKTRAEIQHESGILGPLLYGEVDTL 480  
 421 pddrsysyqlnngpqirgrkkrvpmaytdeftktraeiqhesgilgpllygevdtl 480

QY 421 PDBRSVSYQLNNGPQIRGRKKRVPMAYTDETFKTRAEIQHESGILGPLLYGEVDTL 480  
 421 pddrsysyqlnngpqirgrkkrvpmaytdeftktraeiqhesgilgpllygevdtl 480

QY 481 LIIFKNASRPYNTYPHGITDVRLYRRLPKGVKHLKDPILPGELFKKWTATVEDGP 540  
 481 liifknasrpyntyphgitdvrlrysrlpkgvkhlkdpilpgelfkkwtatvedgp 540

QY 481 LIIFKNASRPYNTYPHGITDVRLYRRLPKGVKHLKDPILPGELFKKWTATVEDGP 540  
 481 liifknasrpyntyphgitdvrlrysrlpkgvkhlkdpilpgelfkkwtatvedgp 540

QY 541 TKSDPRLTRYSSFVMMERDLASGLIGPLLYCESVDQNGNIMSDKRNVLFSVDE 600  
 541 tksdprltryssfvmmerdlasgligpllycesvdqngnimsdkrnvlfsvde 600

QY 541 TKSDPRLTRYSSFVMMERDLASGLIGPLLYCESVDQNGNIMSDKRNVLFSVDE 600  
 541 tksdprltryssfvmmerdlasgligpllycesvdqngnimsdkrnvlfsvde 600

QY 601 NRSWYLTENIQRLPNPAGVQLEDPEFOASNIMHSINGYVDSIQLSVCLHEVAYWYLS 660  
 601 nrswylteniqrlpnnpagvqledpefoasnimhsingyvdslqslshevaywyls 660

QY 601 NRSWYLTENIQRLPNPAGVQLEDPEFOASNIMHSINGYVDSIQLSVCLHEVAYWYLS 660  
 601 nrswylteniqrlpnnpagvqledpefoasnimhsingyvdslqslshevaywyls 660

QY 661 IGAOTDLSVFFSGYMPFKHKVYEDTLTFPFGSEYFEMENENGLMILGCHNSDFRNRG 720  
 661 igatdlsvffsgymfkhkvyedtltfpfgseyfemenenglmilgchnsdfrnrg 720

QY 661 IGAOTDLSVFFSGYMPFKHKVYEDTLTFPFGSEYFEMENENGLMILGCHNSDFRNRG 720  
 661 igatdlsvffsgymfkhkvyedtltfpfgseyfemenenglmilgchnsdfrnrg 720

QY 721 MTALLKVVSSC---DKNTGDY-----BDSYEDIASYLKSNNALEPSSONSHPHS 769  
 721 mtallkvssc---dkntgdy-----bdsyediasylksnnalepssonshpshs 769

QY 721 MTALLKVVSSC---DKNTGDY-----BDSYEDIASYLKSNNALEPSSONSHPHS 769  
 721 mtallkvssc---dkntgdy-----bdsyediasylksnnalepssonshpshs 769

QY 770 TROKQENAT----- 778  
 770 trokqenat----- 778

QY 775 TRGQFNATLIPENDIEKTDQPFHRTCPMKIQNVSSDILMLLGRPTPHGLSLSLDGE 834  
 775 trgqfnatlipendiektdqpfhrtcmpkignvssdilmlmgrptphglslsldge 834

QY 775 TRGQFNATLIPENDIEKTDQPFHRTCPMKIQNVSSDILMLLGRPTPHGLSLSLDGE 834  
 775 trgqfnatlipendiektdqpfhrtcmpkignvssdilmlmgrptphglslsldge 834

QY 835 AKYELTSDPSPAIDSNLSLSEMTHFRPQLHNSGDMVFRPESGILRLNEKLGTTADPL 894  
 835 akyletspdpspaidsnlsllsemthfrpqlhnsghdmvfrpesgilrlneklgttadpl 894

QY 835 AKYELTSDPSPAIDSNLSLSEMTHFRPQLHNSGDMVFRPESGILRLNEKLGTTADPL 894  
 835 akyletspdpspaidsnlsllsemthfrpqlhnsghdmvfrpesgilrlneklgttadpl 894

QY 895 AWDNHYGTQJPKKEWKSQEKSPKTAFKKDDLLSLNACSNHAIASINQNKPELEVT 954  
 895 awdnhygtqjpkkewksqekspktafkkddllslnacsnhaiasinqnkpelevt 954

QY 895 AWDNHYGTQJPKKEWKSQEKSPKTAFKKDDLLSLNACSNHAIASINQNKPELEVT 954  
 895 awdnhygtqjpkkewksqekspktafkkddllslnacsnhaiasinqnkpelevt 954

QY 779 -----PVLAKRHQREITRTTLQSDOEIDYDDTISVEKKKEDFDIYDEDN 824  
 779 -----pvlakrhqreitrttlqsdoeidyddtisvekkkedfdiydedn 824

QY 779 -----PVLAKRHQREITRTTLQSDOEIDYDDTISVEKKKEDFDIYDEDN 824  
 779 -----pvlakrhqreitrttlqsdoeidyddtisvekkkedfdiydedn 824

QY 955 WAKGRTERLCSGPPVLKRNHGREITRTCLISGDEEIDYDDTISVEMKKEDFDIYDEDN 1014  
 955 wakgarterlcsppvlkrnhgreitrtclisgdeeeidyddtisvemkkedfdiydedn 1014

QY 825 QSPRSFOKTRRHVFLAAVERLMDYGMSSSPHVLNRRAQSSGSPQFKKVPQEFDTGSFTQ 884  
 1015 qsprsfoktrrhvflaaverlmdygmssspvhlnrraqsqspqfkkvpqefdtgsftq 1074

QY 885 PLYRGELNEHLGLLPYIRAEVEDNIMVTFRNQASRPYSFSSLSIYEEDROGAEPKRN 944  
 1075 plyrgelnehlgllpyiraevednimvtrfnqasrpysslsisyeedrogaepkrn 1134

QY 945 FVKPNEKTYRPMKQHNHAPRKDEFDCKAAMAFSDVDLEKRVHSGLLGPLVCHTNLNP 1004  
 1135 fvkpnektyrpmkqhnhaprkedefdkamafsdvdlekrvhsgllgplvchtnlnp 1194

QY 1005 AHGRQVTVQEEALFFETJFDEKSWYFTEENMRNCRAPCNIMEPTPEKENYRFAINXI 1064  
 1195 ahgrqvttvqealffetjfdkswyftennmrncrapcnimeptpekentyrfainxi 1254

QY 1065 MDTLPLGLVMAODORIHWYLSMGSNENIHSIHFGHVEYVKKKEEYKALNYLPGVEET 1124  
 1255 mdlplglvmaodorihwylsmgsnenihshfghvetyvkkkeeykalnylpgveet 1314

QY 1125 VEMLPKSAKIRVBCVLIGEHLHAGMSTFLVYSNKCOTPLGMSAGHIRDPOITASGORYO 1184  
 1315 vemlpksakirvbcvligehlhagmstflvysnkcotplgmsaghirdpoitasgoryo 1374

QY 1185 WAPKLARLHYSGSTINAMSTKEPFSWIKVDLAPMIIHGIRKQGRKORFSSIXISOFITMY 1244  
 1375 wapkrlarlhysgstinamstkepfswikvdlapmiihgirkrqgrkorfssixisofitmy 1434

QY 1245 SLDDCKMOTYRGNSTGTLMVFGVNDSSGICKHNFENPPIIARYTRLPHTHSISITLME 1304  
 1375 slddckmotyrgnstgtlmvfgvndssgickhnfenppiiarytrlphthsisitlme 1434

QY 1305 LMGCDLNSCKPCLMESKASISDAQITASSYTFNNMFATWSPSKARLHQGRSNANRPVYN 1364  
 1495 lmgcdlnsckpclmeskasisdaqitassytfnmfatwspskarlhqgrsnanrpvyn 1554

QY 1365 PKEWLOYDFOKTMYVTGVTNGVSKSLTSMVYKEFLISSODGHOQLFPONGKVKVFOG 1424  
 1555 pkewloydfoktmyvtgvtngvsksltsmvykeflissodghoqlfpongkvkvfog 1614

QY 1425 NQDSFTPVNSLDPPLLTRYLRIHQSVWHDIAUMEVLCGEAODY 1471  
 1615 nqdsftpvnsldpplltrylrihqsvwhdiaumevvlcgeaody 1661

RESULT 8  
 AAP80265  
 ID AAP80265 standard; protein; 1516 AA.  
 AC AAP80265;  
 DT 10-OCT-1990 (first entry)  
 XX  
 DE Modified factor VIII:C sequence with the Q744-D1563 deletion.  
 XX  
 KW Modified factor VIII:C; maturation polypeptide; haemophilia;  
 KW blood coagulation; QD deletion.  
 XX  
 OS Homo sapiens.  
 OS  
 PN W08800831-A.  
 XX  
 PD 11-FEB-1988.  
 XX  
 PF 31-JUL-1987; 87WO-US01814.  
 XX  
 PR 01-AUG-1986; 86US-0893375.  
 XX  
 PA (BIOU ) BIOGEN NV (PASE//).  
 XX  
 PI Pasek MP;  
 XX



KW half-life; haemophilia; mutant; mutein.  
XX  
OS Homo sapiens.  
XX  
PN W020071714-A2.  
XX  
PD 30-NOV-2000.  
XX  
PE 24-MAY-2000; 2000WO-US14111.  
XX  
PR 24-MAY-1999; 99US-0135847.  
XX  
PA (AMNA-) AMERICAN NAT RED CROSS.  
XX  
PI Saenko EL, Strickland DK;  
XX  
DR WPI; 2001-025163/03.  
XX  
PT Factor VIII mutants having increased half-life useful for treating  
PT hemophilia, comprise one or more amino acid substitutions in the A2  
PT and/or C2 domain of factor VIII -  
XX  
PS Claim 9; Fig 2A-B; 121pp; English.  
XX  
CC The invention relates to human factor VIII mutants comprising an amino  
CC acid substitution at one or more positions in the A2 domain and/or an  
CC amino acid substitution at one or more positions in the C2 domain.  
CC The invention also encompasses a factor VIII mutant which lacks a B  
CC domain (AAB842). The factor VIII mutants have an increased half-life  
CC in the bloodstream. The A2 domain mutants exhibit reduced LRP-dependent  
CC (receptor-dependent) clearance of factor VIII, while C2 domain mutants  
CC have reduced receptor-independent clearance. The invention also relates  
CC to a method of using RAP (receptor associated protein), a protein which  
CC inhibits LRP (low density lipoprotein related protein), mediated ligand  
CC internalisation, to increase the half-life of factor VIII. The mutant  
CC factor VIII proteins, and nucleotides encoding them, are useful  
CC for treating haemophilia. RAP, LRP-binding RAP mutants or fragments, and  
CC nucleic acids encoding them may also be used in the treatment of  
CC haemophilia. In combination with a mutant factor VIII protein or DNA of  
CC the invention. The invention provides means of increasing the half-life  
CC of factor VIII by reducing its clearance from plasma. The present  
CC sequence represents a mutant mature human factor VIII which lacks a B  
CC domain.  
XX  
SQ Sequence 1424 AA:  
  
Query Match 96.4%; Score 7585; DB 22; Length 1424;  
Best Local Similarity 98.0%; Pred. No. 0;  
Matches 1423; Conservative 0; Mismatches 1; Indels 28; Gaps 1;  
  
QY 20 ATRRYVIGAVELSDVYQSDLGELPVDARPPRYPKSPFNTSVYVKKTLFVEFTVHLEN 79  
Db 1 atrryvigavelsdvwqsdgelpvdarpprypkspfntsvyvkktlfvettchfn 60  
  
QY 80 IAKRPMPMGLIGPTIOAEYVDIVITLKNASHPVSLHAGVSYMKASGSAEYDDQTSQ 139  
Db 1 iakrppmgligptioaeyvditvltknashpvsilhagvsymkasgsaeeyddqtsq 120  
61 iakrppmgligptioaeyvditvltknashpvsilhagvsymkasgsaeeyddqtsq 120  
  
QY 140 REKEDRVFPQSGSHYVVOVLKENGPMASDPLCLTYSLSHVDLVKDLNSGLIALLVCR 199  
Db 121 rekedrvfpqsgshyvvovlkenqpmasdpclcltyslshvdlvkdlnsgliallvcr 180  
  
QY 200 EGSIAKKEKQTQTLKFTLLPAVFDGSKSMHSTKNSLMQDRDASARAPMKHTYNGVYNR 259  
Db 181 egsiakkekqtqtlkftllpavfdgsksmhstknslmqdrdaasarapmkhtvnygvnr 240  
181 egsiakkekqtqtlkftllpavfdgsksmhstknslmqdrdaasarapmkhtvnygvnr 240  
  
QY 260 SLPLGICHRKSYVMHYIGMGTTPREVHSIFLEGHTPLVRNHRQASLEISPTFTLAOTLL 319  
Db 241 slplgichrksyvmhyigmgttpevhnsifleghtplvrnhrqasleisptftlaotll 300  
241 slplgichrksyvmhyigmgttpevhnsifleghtplvrnhrqasleisptftlaotll 300  
  
QY 320 MDLQGFLLFCHSISSHQHDGMEAYVVDSCPEEPQLRMKNNEAEADYDDDLTSEMDVYRF 379  
Db 320 mdlqgfllfchsiSSHQHDGMEAYVVDSCPEEPQLRMKNNEAEADYDDDLTSEMDVYRF 379

Db 301 mdlqgfllfchsiShqhdgmeayvkvdscpeepqlrmkneeadydddlTsemdvrf 360  
  
QY 380 DDDNSPFIQIRSVAKKPKTWVHYIAEEDDMYAPLVIAPDDRKYKSOYLNNGPQIG 439  
Db 361 dddnspfiqirsvakkpktwvhyiaeedwyparlviapddrkyksqylnngpqrig 420  
  
QY 440 RYKRVREMAATDEFTFKREAIOHESGILGLGEVODTLIFKQNASPPYNIYPMGI 499  
Db 421 rkykvrremaatdeftfkreaiohesgilglgevodtllifkqnasppyniypmgi 480  
421 rkykvrremaatdeftfkreaiohesgilglgevodtllifkqnasppyniypmgi 480  
  
QY 500 TDVRPLYSRRLPKGVKHLKDPILPGETFKYKWTWVEDGPTKSDPRLRYSSPVNME 559  
Db 481 tdvrplysrilpkgvkhlkdpilpgetfkykwtwvedgptksdprrlrysspvnm 540  
481 tdvrplysrilpkgvkhlkdpilpgetfkykwtwvedgptksdprrlrysspvnm 540  
  
QY 560 RDLASGLIGPLILICYKESVDORGNOIMSDKRNVLFSFEDENRKSWYLENQRFLPNAG 619  
Db 541 rdlasgligplilicykesvdorgnomsdkrnvlfsfedenrkswylenqrflpnag 600  
541 rdlasgligplilicykesvdorgnomsdkrnvlfsfedenrkswylenqrflpnag 600  
  
QY 620 VOLBDPEFOASNIHNSINGYFDSLOLSCHEVAWYIISGAOTDPLSYFSGYEPKH 679  
Db 601 volbdpefoasnihnsingyfdslolscchevawyilisgaotdplsyfsgyepkh 660  
601 volbdpefoasnihnsingyfdslolscchevawyilisgaotdplsyfsgyepkh 660  
  
QY 680 KMVEDTLTLTFPSGETVFMSEMPGLMILGCHNSDFERNQMTALAKVSCDKMTGDYE 739  
Db 661 kmvedtltltfpsgetvfmsempglmilmilgchnsdfernqmtalakovsckmtgdye 720  
661 kmvedtltltfpsgetvfmsempglmilmilgchnsdfernqmtalakovsckmtgdye 720  
  
QY 740 DSYEDISAYLSKNNALFPRFSQNSRNPSTRQOFNATPVYLKRHREIRTLQSQOE 799  
Db 721 dsyedisaylsknnalfprfsqnsrnpstrqofnatpvylkrhreirtlqsqoe 752  
721 dsyedisaylsknnalfprfsqnsrnpstrqofnatpvylkrhreirtlqsqoe 752  
  
QY 800 EIDDDPTISVMEKEDDIDYEDENQSPRSFQKTRHFIYAVERLMDYSSPHVLRN 859  
Db 753 eidddptisvmeKEDDIDYEDENQSPRSFQKTRHFIYAVERLMDYSSPHVLRN 812  
753 eidddptisvmeKEDDIDYEDENQSPRSFQKTRHFIYAVERLMDYSSPHVLRN 812  
  
QY 860 RAQSGSVPOEFKVVPOEFTDGSFTQPLYRGELNHLGLGPLYIAEVEDNIMVTFRMOAS 919  
Db 813 raqsgsvpoefkvvpoeftdgstfqpilyrgelnhlglgplyiaevdnimvtfrrmoas 872  
813 raqsgsvpoefkvvpoeftdgstfqpilyrgelnhlglgplyiaevdnimvtfrrmoas 872  
  
QY 920 RPYFTYSSLSIEEDQOGAPRKKNFYKPNFTKRYPKVQNHMAPTDEPFCKMAAFSD 979  
Db 873 rpyftysslsieedqogaprkknfykpnftkrypkvqnhmaptdepfckmaafsd 932  
873 rpyftysslsieedqogaprkknfykpnftkrypkvqnhmaptdepfckmaafsd 932  
  
QY 980 VDLEKDVHSGILGPLVLCNTTLPNAGROVTOEFALEFIPDETSMWYTEMNERCR 1039  
Db 933 vlekdvhsgilgplvlcNttlpnagrovtoeFalefipdetSmwytemnercr 992  
933 vlekdvhsgilgplvlcNttlpnagrovtoeFalefipdetSmwytemnercr 992  
  
QY 1040 APCNIQMEDPTFKENYRFHAIINGYIMDTLPGLVAADQRIKMYLLSGNSNENIHSIFSG 1099  
Db 993 apcnIQMEDPTFKENYRFHAIINGYIMdtlpglvaadqrikmYllsgnsnenihSifsg 1052  
993 apcnIQMEDPTFKENYRFHAIINGYIMdtlpglvaadqrikmYllsgnsnenihSifsg 1052  
  
QY 1100 HFTVVRKKEEYKMAVLYNYPGVFETVEMKPSKACIMRWYECILGHLIANGSTLEFLVYSNK 1159  
Db 1053 hftvvrkkeeykmaVlyNypgvfetvEmkpskacimrwYecilghliangstleflvysnk 1112  
1053 hftvvrkkeeykmaVlyNypgvfetvEmkpskacimrwYecilghliangstleflvysnk 1112  
  
QY 1160 CQPLGMAAGHIRDFQITASQYQOMAPKLARLHYSGSINAMWSTKEPFSYKIVLDAPI 1219  
Db 1113 cqplgmasghirdfqitAsqyqomApklarLhysgsinamwstkePfsyKivldaPi 1172  
1113 cqplgmasghirdfqitAsqyqomApklarLhysgsinamwstkePfsyKivldaPi 1172  
  
QY 1220 IHGIRKOGARQFSSLYISOFIIMYSLDGKWKQTYRGNSGTGLTVFFGANDSSGIRKHNIF 1279  
Db 1173 ihgirkogarqfsslyisOfiimYslDgkWKqtyrgnsgtGLtvffgandssgirkhnif 1232  
1173 ihgirkogarqfsslyisOfiimYslDgkWKqtyrgnsgtGLtvffgandssgirkhnif 1232  
  
QY 1280 NPPIIARYIRLHPHYSTRSLRMLMGCDLNSGSMPLGMSKALISAQITASSYFTNMF 1339  
Db 1233 nppiiaryirLhpHystrslRmlmgcdLnsGsmplGmskalisaQitassYftnmf 1292  
1233 nppiiaryirLhpHystrslRmlmgcdLnsGsmplGmskalisaQitassYftnmf 1292  
  
QY 1340 ATWSPSKARLHLQGRSNAAMPQVNNPKEMLOVDQKPMKVTGYVTOGVKSLTSMYKEF 1399  
Db 1293 atwspskarLhlqgrsnaampqVnnpkemLoVdqKpmkvtGyvtOgvkslTsmYkef 1352  
1293 atwspskarLhlqgrsnaampqVnnpkemLoVdqKpmkvtGyvtOgvkslTsmYkef 1352  
  
QY 1400 LISSSDGCHOWTLTFQNGKVKVFOGNDSFTPVVNSLDPLRLRYLRIPQSWVHQIALR 1459  
Db 1353 lissSDgchOWtlTFqngKvkVfognDsfTPvvnsldPlrlrylrIpqswvhQialr 1412  
1353 lissSDgchOWtlTFqngKvkVfognDsfTPvvnsldPlrlrylrIpqswvhQialr 1412



|    |      |              |      |
|----|------|--------------|------|
| QY | 1460 | MEVLGCEAODLY | 1471 |
|    |      |              |      |
| Db | 1413 | mevlqceadly  | 1424 |

RESULT 10  
AAP80267  
ID AAP80267 standard; protein; 1425 AA.

AC AAP80267

DT 10-OCT-1990 (first entry)

DE Modified factor VIII:C sequence with the R740-E1649 deletion.

KW Modified factor VIII:C; maturation polypeptide; haemophilia;

KW blood coagulation; RE deletion.

05 Homo sapiens.

PN W08800831-A.

PD 11-FEB-1988.

PF 31-JUL-1987; 87WO-US01814.

PR 01-AUG-1986; 86US-0893375.

PA (BIOJ ) BIOGEN NV (PASE/).

PI Pasek MP;

DR WPI; 1988-049866/07.

DR N-PSDB; AAN80446

PT New DNA sequences encoding modified factor VIII:C - with deletion of DNA encoding maturation polypeptide, useful for high yield transformation.

PS Claim 3; Page 57-58-59-60; 97pp; English.

CC The entire sequence encoding the maturation polypeptide of

CC The full length Factor VIII:C CDNA has two changes with respect to the  
CC published sequence (EPO application 160457):  
CC

CC (Phe to Leu). The product is produced in approx. 20 times higher

CC easily purified. The peptide is used for treating haemophilia A, both

CC See also AAN80444 and AAN80447.

**SQ Sequence 1425 AA;**

Query Match

Matches 1421; Conservative 0; Mismatches 3; Indels 28; Gaps 1;

|    |     |   |     |
|----|-----|---|-----|
| QY | 20  | ATRRYVYGAVELSMDYQMSDGLGELPVARFPPRVKSEFPNTSVYVKKTLFVEFTVHLN    | 79  |
| Db | 2   | atrryylgavelsmdymgsdglgelpvarfpprvkspfntsvvykkelfveftchlfn    | 61  |
| QY | 80  | IAKRPRPMWGLGPTDAEYVDIVYITLKNMASHPSVLAAVGYSTYKASGAEDDOTSO      | 139 |
| Db | 62  | iakrprpwmgligptlgaeydvtvlclknmasphvslhavgvsvywkasegaeyddqts   | 121 |
| QY | 140 | REKDDKVFPGCSHTYVWOYLKEKGPASDPDLCTIYSYLSHVDLVKQNLNGLLGALLVCR   | 199 |
| Db | 122 | rekeddkvfygqsthtyvwqylkengpasdpdlctlsyylshvdlvklngllgallvcr   | 181 |
| QY | 200 | EGSLAKERTQTLHKFTLLFAVFEDEGKSMHSETKNSLMQDDDAASARAMPKMTHTVNGVNR | 259 |

|    |      |   |      |
|----|------|---|------|
| Dh | 182  | esjslkehcthtlhnkfiillfavfdeogkshsetknsimgotdaasaraawpkhmtvngynr | 241  |
| Qy | 260  | slrplglghnksyvmwvniwgmttrebvhslfeghtflvrnhnroasleisptflaotll    | 319  |
| Dh | 242  | slrplglghnksyvmwvniwgmttrebvhslfeghtflvrnhnroasleisptflaotll    | 301  |
| Qy | 320  | mluogflrfchssnqhnogmavaykvvscrebepolrkknnebeedyddoltosemdvyrf   | 379  |
| Dh | 302  | mluogflrfchssnqhnogmavaykvvscrebepolrkknnebeedyddoltosemdvyrf   | 361  |
| Qy | 380  | ddnnspsftoirsvaykknprtvmwvniAAEEEDMDVAPLVARPDRSYKQYLNNGFORIG    | 439  |
| Dh | 362  | ddnnspsftoirsvaykknprtvmwvniAAEEEDMDVAPLVARPDRSYKQYLNNGFORIG    | 421  |
| Qy | 440  | RYKKVRFMAVYTDDETFKTREAIOHESGILPRLYGEVGDTLIIFFKNOASREPNYIPIGI    | 499  |
| Dh | 422  | RYKKVRFMAVYTDDETFKTREAIOHESGILPRLYGEVGDTLIIFFKNOASREPNYIPIGI    | 481  |
| Qy | 500  | PNVRPLYSRLRPKYGNKHKDPRILPGEIPIKYKWTYVEDOPRYSDBRCILTRYSSVAME     | 559  |
| Dh | 482  | PNVRPLYSRLRPKYGNKHKDPRILPGEIPIKYKWTYVEDOPRYSDBRCILTRYSSVAME     | 541  |
| Qy | 560  | RLAAGLGLPRLICYKESVDORGNQMSDRNVIILSVPEPNSWVLTENIORPLPNAG         | 619  |
| Dh | 542  | RLAAGLGLPRLICYKESVDORGNQMSDRNVIILSVPEPNSWVLTENIORPLPNAG         | 601  |
| Qy | 620  | VOLDEPERQASINMHSINGVYFDSLQSLVCLNHEVAYWILISTGAQDFLVSFSGTEKH      | 679  |
| Dh | 602  | VOLDEPERQASINMHSINGVYFDSLQSLVCLNHEVAYWILISTGAQDFLVSFSGTEKH      | 661  |
| Qy | 680  | KWVYEDTLLEPFSGEYVFNMSMNPGLMIIGCHNSDFRNEMTALLKVSCKNTEGYE         | 739  |
| Dh | 662  | KWVYEDTLLEPFSGEYVFNMSMNPGLMIIGCHNSDFRNEMTALLKVSCKNTEGYE         | 721  |
| Qy | 740  | DSYEDISAYLXLSKNNALTEPRRSFSONSNHPRSTROKOFNAPRVLKRNQREITPTTQDSOE  | 799  |
| Dh | 722  | DSYEDISAYLXLSKNNALTEPRRSFSONSNHPRSTROKOFNAPRVLKRNQREITPTTQDSOE  | 753  |
| Qy | 800  | EIDYDDTLTSVEMKKEDEPIYDEDENOSPFSFOKTRHFAVERLMDVGMSSPHVLYRN       | 859  |
| Dh | 754  | EIDYDDTLTSVEMKKEDEPIYDEDENOSPFSFOKTRHFAVERLMDVGMSSPHVLYRN       | 813  |
| Qy | 860  | RAOGSVPOFKVVPFOETDGSFTQPLRYGELNENILGLBRYTRAEBEDNIWYTRNOAS       | 919  |
| Dh | 814  | RAOGSVPOFKVVPFOETDGSFTQPLRYGELNENILGLBRYTRAEBEDNIWYTRNOAS       | 873  |
| Qy | 920  | PRYSYSSLISZVERDOSGABRRKNAPUKREKTEKTYRMKONHNAAPRKDEFOCKAAAYUSD   | 979  |
| Dh | 874  | PRYSYSSLISZVERDOSGABRRKNAPUKREKTEKTYRMKONHNAAPRKDEFOCKAAAYUSD   | 933  |
| Qy | 980  | VOLEKDVHSGILGVLACHNTLNPAPRAGROVYUQOEALFTTIFDETKSMVYFENNEMRRCR   | 1038 |
| Dh | 934  | VOLEKDVHSGILGVLACHNTLNPAPRAGROVYUQOEALFTTIFDETKSMVYFENNEMRRCR   | 993  |
| Qy | 1040 | APCNIOMEDPTFKENRYRPHAINGYIMDTPLGVLMAODIRIWMYLLSMGSENIHISIRHSG   | 1099 |
| Dh | 994  | APCNIOMEDPTFKENRYRPHAINGYIMDTPLGVLMAODIRIWMYLLSMGSENIHISIRHSG   | 1053 |
| Qy | 1100 | HVFYTRKKEEYKALNLYLRYEVEYVYMLRBSKAGIRVBSCLIGENHAGMSTLFLVYLSNK    | 1155 |
| Dh | 1054 | HVFYTRKKEEYKALNLYLRYEVEYVYMLRBSKAGIRVBSCLIGENHAGMSTLFLVYLSNK    | 1113 |
| Qy | 1160 | COTPLGMAVSGHIRFOFOTIASOYGOVAPKTLARLHSGSINAMSTKPRPFMIIVDLALMI    | 1219 |
| Dh | 1114 | COTPLGMAVSGHIRFOFOTIASOYGOVAPKTLARLHSGSINAMSTKPRPFMIIVDLALMI    | 1173 |
| Qy | 1220 | IHGIKTOGAROFESLSLTSOFLIIMYSLDGGKMMQTYRGNSGTGLMVFEGVNOSSGIKHNIIF | 1279 |
| Dh | 1174 | IHGIKTOGAROFESLSLTSOFLIIMYSLDGGKMMQTYRGNSGTGLMVFEGVNOSSGIKHNIIF | 1233 |
| Qy | 1280 | NBPITARIYRLHPTHYISIRSTLARMELMGCDLNSCMPRLMESKASISDAOITASSYFTYMF  | 1339 |
| Dh | 1234 | NBPITARIYRLHPTHYISIRSTLARMELMGCDLNSCMPRLMESKASISDAOITASSYFTYMF  | 1293 |



Db 1113 cqlplmasghlrdqfqlasggygwapklarlhysgsinawskpefswikvdlapml 1172  
 Qy 1220 IHGIRTOGAROKFSSLYISOFTIMYSLODKKWOTYRGNSTGTLMVFEFVNDSSGIKHNIF 1279  
 Db 1173 ihgltqgarqcfsslylsqfllimysldgkkwqlyrgnstgltlmvffgnvdssgikhnlf 1232  
 Qy 1280 NPPIIARYIRLHPHYSTIRSTRLMELMGCDLNSCSNPLMGESKAISDAOITASSSYFTNMF 1339  
 Db 1233 npplliarylrhlphlysrstlrmlmgcdlnscsnplmgeskalsdaqflassyfllnmf 1292  
 Qy 1340 ATWSPSKARLHOGRSNMNRPOVNNPKELQYDFOKTMKVTGVTTOGVKSLTSMYKKEF 1399  
 Db 1293 atwspskarlhlgfnsnawrpvnnpkewlgydfoktmkvtygtvgvkslltsmykcef 1352  
 Qy 1400 LISSQDGHOWTLFPONGKVKVFOGNOGSFTPVNSLDPPLLTRYLRIHPQSVWQIALR 1459  
 Db 1353 lissqdghowtlfpgnkvkvfognosftpvnsldpplltrylrihposwvhoialr 1412  
 Qy 1460 MEVLGCEADLY 1471  
 Db 1413 mevlgceadly 1424  
 RESULT 12  
 AAP91169  
 ID AAP91169 standard; protein: 1424 AA.  
 AC AAP91169;  
 XX  
 DT 26-JUN-1990 (first entry)  
 DE Sequence of 740 Arg-1649 Glu human Factor VIII:C.  
 KM Human Factor VIII:C; Ad.RE.neo: 740 Arg-1649 Glu Factor VIII:C;  
 KW haemophilila A.  
 OS Homo sapiens.  
 PN EP306968-A.  
 PD 15-MAR-1989.  
 PE 09-SEP-1988; 88EP-0114769.  
 PR 08-APR-1988; 88JP-0085454.  
 PA (KAGA ) CHEMO-SERO-THERAP (TEIJ).  
 PI Sugiyama T, Masuda K, Tajima Y, Yonemura H;  
 DR WP1: 1989-078467/11.  
 DR N-PSDB; AAN90654.  
 PT Prodn. of recombinant human Factor-VIII-C -  
 PT using animal cells transformed with a vector contg. the gene for  
 PT Factor VIII:C and a promoter  
 PS Fig 1(1) - 1(13); : 32pp; English.  
 XX Arg-740 of the carboxyl terminus of the H chain is  
 CC directly bonded by a peptide bond to Glu-1649 of the amino terminus of  
 CC L chain. A prefd. expression vector used to transform animal cell so  
 CC that they produce human Factor VIII:Cis plasmid Ad.RE.neo.  
 CC The expression vector has at least one promoter upstream of AAN90654.  
 CC The transformants can constantly and continuously produce human Factor  
 CC VIII:C in high yield on a commercial scale. The human Factor VIII:C so  
 CC produced is considered to corresp. to the smallest species of active and  
 CC intact Factor VIII:C molecules in the human blood plasma. It is useful  
 CC for treating haemophilila A patients.  
 SO Sequence 1424 AA:

Query Match 96.2%; Score 7567; DB 10; Length 1424;  
 Best Local Similarity 97.8%; Pred. No. 0;  
 Matches 1420; Conservative 0; Mismatches 4; Indels 28; Gaps 1;  
 Qy 20 ATRRYTLGAVELSWDMQSDLELYDARFPPRVKSPSPFNISVYKKTLYVEFVHLFN 79  
 Db 1 atrrytlgavelswdmqsdlqelpvdarfprrvksfpfnisvyykktllyvefdhlfh 60  
 Qy 80 IAKRPWMGLLQPTIOAEVYDTVYITLKNMASHVSNLHAGVSWKSEGAEDDQTSQ 139  
 Db 61 iakrpwmgllypqltgaevydtvlytlknmashvsnlhagvswkasegaedqdtsg 120  
 Qy 140 REKEDKVEFGSSHTYVWQVLKENGPMASDPLCLTYSLSHVDLYKDLNSGLIALLVCR 199  
 Db 121 rekedkvefgsshtcywqvlkengmasdplcltyslyshvdlvkdlnsgllgallvcr 180  
 Qy 200 EGSIAKKEKQTLHKFTLLFAVPEDEKSMHSEKNSLMODRDAASRAPKMHYNGVYNR 259  
 Db 181 egslakekthlkhftllfavpdegkswhselknsldmdraasarapkmhlyvnyvr 240  
 Qy 260 SLPLGLICHRKSVYMHVIGMTTPVHSTFLEGHFVLNRHROASLESTPTFLTAOTIL 319  
 Db 241 slpllgichrksvywhvigmgttpevhsfliegthflvrnhrgaslelspfltaqltl 300  
 Qy 320 MDLGOFLFCHTISSHQHDMEAIVKVDSCPEEPQLRMKNNEAEDYDDLTDEMDVVR 379  
 Db 301 mdlgqflfchissqhdmeyvkvdscopeepqlrmkneeaedyddltdeemdvr 360  
 Qy 380 DDNSPSFTQISVAKKHKPTVWHYIAEEDMDVAPLVLPDDRSYKSOYLNNCPQRIQ 439  
 Db 361 ddnspsftqlsvakkhkptvwhylaeeedwdyaplvlpdhdsvksqylmnpqr 420  
 Qy 440 RYKRVKFMAYDETFKTRREALIOHESGLPLLYGEVDTLLIIFKNASPPVNIYHGI 499  
 Db 421 rkykvrkfmaydeltfktreaqhsgllgpllygevgdtlllflknasppvniybhgi 480  
 Qy 500 TVVRPLYSRRLPKGVKHLKDEPILPGEIFKRYKWTVJEDGPTKSDPCLTRYSSFYVME 559  
 Db 481 tvvrplysrllpkgvkhlkdpilpgeifkykwetvtedgptksdpccltryssfyvme 540  
 Qy 560 RDLASGLIGPLLICKESSVDQGNQIMSDKRVILFSVENDRSWYLTENQRLPMPAG 619  
 Db 541 rdlasgllgpllllcyesvdqgnqimsdkrvilfsyvdendrswyltenqrlpmpag 600  
 Qy 620 VOLEDEPEFQASNIMHSINGYVDSLOLSVCLHEVAVWYILSIGAQTDFLSVFSGYFKH 679  
 Db 601 vledpefefqasnimsingyvdslqsvclheavaywylsigqtdlsvsfsgyfkfh 660  
 Qy 680 KMYVEDTTLTLPFSGEFTVPMKSNENPGLWILGCHNSDFRNKGMALILKYSSCDKNTGDYE 739  
 Db 661 kmvyedtltlfpfsgetvfmsnengplwllgchnsdfnrgmtalilkysscdkntgdye 720  
 Qy 740 DSEYEDISAVILSKNNAIPRSPSONSRHPSRQKFNATPVILRHQEIERTLOSDOE 799  
 Db 721 dseyedisavilsknnaip-----retrtlqsdqe 752  
 Qy 800 EIDYDQTSIVEAKKEDFDIYDEENOSPRSFQKTRHYFIAVERLMDYGMSSSPHYLRN 859  
 Db 753 eidydqtsiveakkedfdiydedensprsfqktrhyfiaaverlwdygmsssphylrn 812  
 Qy 860 RAQSSVPOFKKVVPOEFTDGSFTOPLYRGELINEHLGLGPLYIRAEVEDNIMVTFRNQAS 919  
 Db 813 raqssvpofkkvvpoeftdgstplyrgelnehllgplyiraevednmvtfrrnqas 872  
 Qy 920 RYRSFYSSLISYEEDROGAERKRFVFNPKTYFKWKVQHMMAPTKEPFCKAMAYFSD 979  
 Db 873 rrsfysslysyeeqrqgaeprrkfvpnekyfkwvqhmmaptkepfckamayfsd 932  
 Qy 980 VLEKDVHSGILGPLLVCHTNTLPAHAGROYTVOEFAFLFTFIDETKSWYTTENMERCR 1039  
 Db 933 vlekdvhsgllgpllvchtntlnpahagroytvefalfftidetkswytememcr 992  
 Qy 1040 APCNTIOMEDPTEKENYRPHAINGYIMDTLPGLVMAQDQIRIMWILLSMGSENHISHIFSG 1099

|||||  
Db 993 apcnlmedpfkenyrfhalingymdclpglmaeqdgrirwyllsmgsneenlhahtesg 1052  
QY 1100 HVEFRRKEEKEMALYNLYPGVEFEVEMLPKAGIWRRECLIGELHMGMSLFLYVSNK 1159  
Db 1053 hvflvrkkeeYkmalayllypvfclvemlpkaglwrecllgehlhagmstlflvysnk 1112  
QY 1160 COTPLGMSGHI RDEQITASCOYGOMAPRLARLHYSGSISINMWSTKEPFSWIKVDLLAPMI 1219  
Db 1113 cqlptgmasghlrdtqtasqyggwapkrlarlhysgsinawstkeptswlkvdllaemi 1172  
QY 1220 IHGKTGAKOKFESSLYISQPTIMVSLDGKKQVYRGNSGTGLMWFEGNVDSGIRHNI 1279  
Db 1173 lhgltggarqkfsslylsqglfimsldgkkwqyrgnstgltlmwffgnvdsqgkhnlf 1232  
QY 1280 NPPIIARIIRHPTHYSTRSLRMLMKCDLNSCSMPRGMSKASIDQITASSFTWMP 1339  
Db 1233 nppliarylrlhpthtystrslrmlmcdlnscsmplrgmskalsdaqitassyltmf 1292  
QY 1340 ATWSPSKARLHLOGRSNAMPQVNNPKEMLOVDFOKTKMTGVTTOGYKSLTSMYKKEF 1399  
Db 1293 atwspskarlhllogrsnawrpqvnmpkewlgydfgkcmkvgttgqkslltmyvkef 1352  
QY 1400 LISSSDGHQMTLEFFQNGKVKVFOGNDSEFPVYNSLDPPLLTRYLRILHPOSWVHQIALR 1459  
Db 1353 lissqdgqhtlffgnqkvkvfgngdsfcpvynslpplltrylrhlpqswvhqialr 1412  
QY 1460 MEVLGCEADIX 1471  
Db 1413 mevlqceadily 1424

## RESULT 13

AAW11422 standard; Protein: 2342 AA.

AAW11422:

20-NOV-1997 (first entry)

Active Factor VIII:C analogue, delta 1311-1320, + Pro insertion.

Factor VIII:C; analogue: glycoprotein; blood coagulation cascade;  
fibrinogen; fibrin clot; haemostasis: haemophilia A; bleeding diathesis;  
plasma protease; thrombin; immunogen: antibody; haemophilic; therapy;  
proteolytic cleavage.

Homo sapiens.  
OS  
Synthetic.

Location/Qualifiers

1..19

Peptide

/note= "signal peptide"

20..2342

Protein

/note= "mature Factor VIII:C"

20..1658

Region

/note= "heavy chain fragment"

Misc-difference 1329..1330

/note= "site of 10 residue deletion"

Misc-difference 1330

/note= "inserted residue"

Region

1659..2341

FT

/note= "light chain fragment"

Domain

760..1658

/note= "B domain"

XX

XX

XX

XX

XX (CHIR ) CHIRON CORP.  
XX  
XX

PI Cohen FE, Hung DT, Innis M;  
XX  
XX

DR WPI: 1997-119050/11.  
XX

PT Factor VIII:C analog modified adjacent to a non-activating Arg  
PT residue - used in the treatment of haemophilias, by improvement of  
PT haemostasis  
XX

XX Claim 27: Page -: 90pp: English.

AAW11330-W11472 represent active Factor VIII:C analogues of the  
invention. These sequences were created by mutating the wild type Factor  
VIII:C coding sequence (see AAT51357) using mutagenic primers. The  
analogues comprise a native Factor VIII:C polypeptide modified at a site  
adjacent to a non-activating Arg residue so that a Arg-Pro or Pro-Arg  
dipeptide is created. Factor VIII:C is a large glycoprotein that  
participates in the blood coagulation cascade that ultimately converts  
soluble fibrinogen to insoluble fibrin clot, effecting haemostasis. A  
deficiency in Factor VIII:C is responsible for haemophilia A, which is an  
X-chromosome-linked inherited bleeding diathesis. Factor VIII:C is  
activated by plasma proteases, such as thrombin. During activation the  
mature polypeptide is cleaved to generate heavy and light chain fragments  
that are further cleaved. Complexes of two or more of the analogues,  
nucleic acids and vectors encoding them may be used alone or in  
conjunction with each other, for the prevention or treatment of active  
Factor VIII:C deficiency in a mammal. The analogues may be used as  
immunogens to raise antibodies, and in the treatment of haemophilias, by  
improvement of haemostasis. The analogues are resistant to proteolytic  
cleavage and display increased plasma half-life. They may be administered  
at lower dosages and by different modes of administration.

Sequence 2342 AA:

Query Match 94.2%; Score 7416.5; DB 18; Length 2342;

Best Local Similarity 62.8%; Pred. No. 0; Matches 1470; Conservative 0; Mismatches 1; Indels 871; Gaps 1;

QY 1 MQIELSTCFEFLCLRFCSATRRRYLAGAVELSDWQSDLGELPYDARFPFPPKSPFPN 60  
Db 1 mqlstcfllclrfcsatrrylagavelsdwqsdlgelpydarfpfpvpskspfn 60  
QY 61 TSVYKKTLEVEFTVHLFNIAKRPPEWMLGPTIQAEVDTVTITKNMASHVSLHAV 120  
Db 61 tsvykktlfeftdhlfniakrppwmlgptiqaevydtvtitknmashvslhav 120  
QY 121 GVSWKASEGAEYDDQTSQREKEDKVPFGSGSHRYVMQVLKENGPMASDPLCLITYSLH 180  
Db 121 gvsywkasegaeyddqtsqrekedkvfpfgsgshryvmqvlkengpmasdpclclityslh 180  
QY 181 VDLVKDLSGLIGALLVREGSLAKETQTLHKFTLLFAVDEGKSWSEFNKNSLMQORD 240  
Db 181 vdlvkdlsgllgallvregslaketqtlhkftllfavdegkswsefnknslmqord 240  
QY 241 AASARAWPKMHTVNGVYNSRLPGLIGCHRSKSVYWHVIGMGTPEVHSIFLGGHTFLVRNH 300  
Db 241 aasarawpkmhtvngvyvnsrlpgligchrsksvywhvigmgtpevhsiflgghtflvrnh 300  
QY 301 ROASLEISPIETFIPTAOTLLMDLGFLLFCHTSSHOHDMGAVYKVDSCPEEPQIRMKNNE 360  
Db 301 roasleispietfiptaotllmdlglfllfchtsshohtmgaavykvdsceppeqirmkne 360  
QY 361 EAEDYDDDLTDSQEMDVVRFDDNSPFIQIRSVAKKHKPKTWHVYIAAEEDMDVAPLVIA 420  
Db 361 eaedydddltdsqemdvvrfddnspsfiqirsvakkhpkwtwhvyaeeedmdvayplvia 420  
QY 421 PDDRSYKSQYILNNGPQIRGRYKKVREMAVYDEFRTKTRERAIQIHESGILGPLLIGEVDTL 480  
Db 421 pddrsyksqyilnngpqirgryrkkrvremavynetftrtktreraiqihesgilgplligevgdtl 480

QY 481 LIIRKNOASRPNTNPHGTTDVRPLYSRRLPKGVKHLKDFPLLPGEITFKYKMTVYVEDCP 540  
DB 481 IIRKNGESRPYnIYPHgltdvrplysrrlPKYVKhLkdfllpgeIfkyKwvtvvedcp 540  
QY 541 TKSDPRCLTRYSSFPVNNERDLASGLIGPLICVKEVSDORNOIMSDRNVILFESVDE 600  
DB 541 tkspdrcltryssfvmerdlasglIgpIlIcYkesvdqrngqmsdkrrnvlIsvfve 600  
QY 601 NRSWYLENIORFLPNPAGVQLEDEPEQASNIMHSINGVFDLSLQVCLHEVAYWYILS 660  
DB 601 nrswyItenIqrfllpnpagvqledpefgasnimhsingvfdslqsvclheavaywyls 660  
QY 661 IGAOTDFLSVFFSGSTFKHKMYEDTLTLFFPSGTFVFMNMNPGMLIGCINSDFRNG 720  
DB 661 IgaotdfIsvffsgytfkHkmyedtltlffpsgTfvfmnmnpglwlIgcHnsdfrrng 720  
QY 721 MTALIKVSSCDKNFGDYEDSYEDISAYLLSKNNAIEPFSFONSHPSTROKOFNAT-- 778  
DB 721 mtaIlIkvsscdknfgdyedsyedIsayllsknaIeprsfnsrhpsttrqkqfnatcl 780  
QY 779 ----- 778  
DB 781 pendiektcpwfahrtcpmpkIqnvassdIlmlIqspcpHglSlsdlgeakvetsddps 840  
QY 779 ----- 778  
DB 841 pgaIdsmnslsemthfrpqlIhsgdmvftpesqlrlInekIgtateleklldfkvsst 900  
QY 779 ----- 778  
DB 901 snuIlstIpsdnlaagtntslIgpssmpvhydsqldtclfgkkspsltesgplslsee 960  
QY 779 ----- 778  
DB 961 ndskllIesglmsqesagkvustesgrlfkgrahgpalltkdnalfkvsIglIken 1020  
QY 779 ----- 778  
DB 1021 ktsnusatnrthIdpsalIleenspvguIllesdtefkvtrplIhdrlmldknatalrI 1080  
QY 779 ----- 778  
DB 1081 nhmsnktsskrmemvqkqkegplppdaqndmsffkmlfIpesarwIqtrngkslnsg 1140  
QY 779 ----- 778  
DB 1141 qgspkqIvslgpeksveggnfIseknkvvgkgeftkdvgIkemvfpesrnlfltnldh 1200  
QY 779 ----- 778  
DB 1201 lhennthnqekkgIgeelekeclIqenvnlpqIhtvtgtnfkmnlflIstrqnevsgyd 1260  
QY 779 ----- 778  
DB 1261 gavyavIqdfslndstrnrkhtahfskgeenleaglmgtkqIvekyacltrispnt 1320  
QY 779 ----- 778  
DB 1321 sqgnfvtrpplleeteleklrlIvddstfswsknmhltpscltIqIdynekekgaltqpl 1380  
QY 779 ----- 778  
DB 1381 sdcItrshslpqnrsplIakvsfIsrlrplvlfqdnshlpaasyrkkksyqge 1440  
QY 779 ----- 778  
DB 1441 sshIlgagaknnIsalatllemtgdqrevsgIgtatsnsvlykKvenvlpKpdlpkcsq 1500  
QY 779 ----- 778  
DB 1501 kvellpkvhyqkdIfpetsngspghIdIvegsllqglegakIvneanrpqKvflIra 1560  
QY 779 ----- 778

DB 1561 tessaktpskllldplawdnhygtIpkEeWskgspektalfkkdclIsInacesnhaI 1620  
QY 779 ----- PVLKXHOEIRTRTLQSDDEIDYDDTISV 809  
DB 1621 aalIneqmkpeIevetwakqgrterIcosqnpvIkrhqrgeIrtltIqsgedIdydtclsv 1660  
QY 810 EKKKEDFDIYDEDENOSPRSFOKTRRHVFIAAVERLMDYGMSSSPHVLRNRAQSGSVPOF 869  
DB 1681 emkKedfdIdydedengsprsfqkklrhyfIaaverlwdygmssphvIrnraqsgsvpqf 1740  
QY 870 KKVVOEFNDGSEFQPIVARGELNEHLGLGPYIRAEVEDNIMVFRNOSAPRYSFSSLI 929  
DB 1741 kkvvtIqetdgsfIcqpIyrgelnehgIlIgpYIRAEvednImvtrtnqasrpyssll 1800  
QY 930 SYEEDROGAEPKRNKFNVPNETKTYFWYVOHMAPTKDEFDCKAMAYSDVLEKDVHSG 989  
DB 1801 syeedqrGaepkrhfvPnektkyfkwvqhmaprkefckekawysdvdlekdvhsg 1860  
QY 990 LIGPLVCHTNTLNPAGHGRQVTVQEFALFETIPEDKSWYFTEMMERNCRAPCNIQMEDP 1049  
DB 1861 lIgpIlVchntlnpahgrqvvtvgefaIfclfdetksyftemerncrapcnIqmedp 1920  
QY 1050 TEKENYRPHATNGYIMDTLPGLVMAQDORIRNYILLSMGSNENIHSIHPSGHVFTVRKKEE 1109  
DB 1921 tfeKenyrfhaingyImdtlpglvmagdqrlrwyIlsmgsnenIhsIhsgHvftvrkkee 1960  
QY 1110 YKMALYNYLPGVFETVEMLPKAGIWRVECLIGEHLHAGMSTLFLVYSNKCOTPLGNASG 1169  
DB 1961 ykmalYnlypvyfelveMlpkagIwrveclIgehlhagmstlflvysnkcotplgnasg 2040  
QY 1170 HIRFOQTASGOYGQWAPKLARLHAYSGSINAMSTKEPSSWIKVDLAPMIHGIKTGAR 1229  
DB 2041 hIrdfqItasgqygqwapklarlhysgsInawstkePsswIkvdlIapmIhgiKtggar 2100  
QY 1230 OKFSSLYXSOFITIMYSLDGKKWOTRGNSGTGMVFRPNVSSGIKHIEPPIIARVIR 1289  
DB 2101 qkfsslysqfIlImyslIqgkKwqlyrgnsgtlmVfIqnvadsgIkhnlfrpIlaIrylr 2160  
QY 1290 LHPHTYSIRSTLRMELMGCDLNSCSPMLGMSKASISDQIIRASSFTYIMFAWSPSKARL 1349  
DB 2161 lhphtysIrstlrmeImgcdlnscsmpIgmekalsdsqIcassyfcmfctwpskarl 2220  
QY 1350 HUGRSNAMRPOVNNPKEMLOVFOKTMKVTGVTTOGYKSLTSMYKEPLIASSODGHO 1409  
DB 2221 hIqgrsnaWrPvnnpkewlgydfqtkmkvgvtctgykkslltsmyvkeflIsssgdgh 2280  
QY 1410 WTLFFQNGKVKYFGCNOGNSFPVNSLDRPLLTRIRLRHPQSWHQTALRHEVIGCEAOD 1469  
DB 2281 wclffqngkvkvIqngqdsfIprvnsIdrpIlItrylrIhpswvhqalrmevIvigeaqd 2340  
QY 1470 IY 1471  
DB 2341 IY 2342

RESULT 14  
AAW10591  
ID AAW10591 standard; protein: 2351 AA.  
XX  
AAW10591;  
AC  
XX  
DT 03-DEC-1997 (first entry)  
XX  
DE Factor VIII:C (Phe652His).  
XX  
XX Factor VIII:C; F8C; Factor V A; Factor V C; domain; F8C deficiency;  
KM haemophilia A; blood clotting disorder; immunogen; antibody.  
OS  
XX Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..19

FT /note="Signal peptide"  
FT Protein 20..2351  
FT /note="Mature factor VIII:C"  
FT Misc-difference 671  
FT /label= phe652His  
XX  
XX MO9703191-A1.  
XX  
XX 30-JAN-1997.  
XX  
XX 28-JUN-1996; 96WO-US11013.  
XX  
XX 11-JUL-1995; 95US-0001030.  
XX  
XX (CHIR ) CHIRON CORP.  
XX  
XX Hung DT;  
XX  
XX WPI; 1997-119047/11.  
XX  
XX Factor VIII:C analogue - modified to comprise Factor V A or C domain  
XX or subdomain, for increased stability or activity  
XX  
XX  
XX Claim 6; Page -: 45pp; English.  
XX  
XX The sequences given in W10590-92 represent active Factor VIII:C (F8C)  
XX C polypeptide analogues. These analogues comprises a native F8C  
XX C polypeptide modified to comprise the presence of a Factor V A and/or  
XX C domain and/or subdomain. The F8C polypeptide analogues, alone or in  
XX combination, can be used for the prevention or treatment of an active  
XX F8C deficiency, i.e. haemophilia A and other blood clotting disorders.  
XX The analogue can also be used as an immunogen for antibody production.  
XX C The analogues can have an increased plasma half-life, or specific  
XX C activity. This sequence is not given in the specification and is  
XX based on the sequence derived from Genbank ref. K01740.  
XX  
XX  
XX Sequence 2351 AA:  
SQ  
Query Match 94.28; Score 7416; DB 18; Length 2351;  
Best Local Similarity 62.5%; Pred. No. 0;  
Matches 1470; Conservative 1; Mismatches 0; Indels 880; Gaps 1;  
QY 1 MOELSTCFCLRFCEFSATRRYYLGAVELSWDMQSDGELPVDARFPRVPKSPFN 60  
DB 1 mqlstcfcllrfcefsatrryylgavelswdymgsdlgelpvdarfprrvpksfpln 60  
QY 61 TSVYKKTLEVEFTVHLFNIAKPRPMGLGPTIOAEVDTVVITLKNNASHPVSLHAV 120  
DB 61 tsvykktlfevftvhlfnlaktprpmmglgptioaevdttvittknnashpvslhav 120  
QY 121 GVSYWKASGEAEYNDQTSOREKEDKVPFGSGSHYVNOVLKENGPMASDPLCLTYSLSH 180  
DB 121 gvsywkasegaeayndqtsorekedkvfpqsgshyvnovlkenpmsdplcltysylsh 180  
QY 181 VDLAKDINSGLIGALLVCRGSLAKEKQTLHKFILLFAFDEKSWHSEKNSIMODRD 240  
DB 181 vdlakdinsgligallvcregslakektqlhkfillfafdegkswhsetksimodrd 240  
QY 241 AASARAPKMHYNGVYVNSLPGILGCHRSVYWHVIGMGTPEVHSIFLEGHTFLVRNH 300  
DB 241 aasarpkmhnygvynsrslpgilgchrsvywhvigmgtpevhsifleghftlvrn 300  
QY 301 RQASLESPITFLAQTLLMDLGOFLLFCHISQHDGMEAYVNVDSCEPEPQLRMKNNE 360  
DB 301 rqaesletpitflaqtllmdlgfllfchissqhdmeyavnvdscepepqlrmkne 360  
QY 361 EAEDYDDLFDSEMDYVRFDDDNPSFIQIRSVAKKHPKTMVHIIAEEDMDVAPLVLA 420  
DB 361 eaedyddldfsemdyvrfdndnpsfiqirsvakkhpktmvvhiaeedmdvaplvl 420  
QY 421 PDDRSYKSOYLNNGPQIRIGRKYKVRPMAYTDETFKTREAIQHSGLIGPLLYGEVGD 480  
DB 421 pddrsykylnngpqirigrkykvrpmaytdefktrreaiohsgilgpllygevgd 480

DB 421 pddrsykylnngpqirigrkykvrpmaytdefktrreaiohsgilgpllygevgd 480  
QY 481 LIIFKNQASREYNITPYGIDTVRRPLYSRRLPKGVKHLKDPFILLGELTFKRYWTVVEGP 540  
DB 481 liifknqasreynitpygidtvrrplysrllpkgvkhlkdpflllgeltfkrywttvegp 540  
QY 541 TRSDPCLTRYSSSFVNMNERDLASGLIGPLICYKESVDQGNQIMSDRRVILFSVDE 600  
DB 541 trsdpcltrysssfvnmnerdlasgligplicykesvdqgnqimsdrrvilsfvde 600  
QY 601 NRSWYLTENIQRLPNPAGVQLEDPEFOASNIMHSINGYVDSQLSVCLHEVAYWYLS 660  
DB 601 nrswylteniqrlpnpagvqledpefoasnimhsingyvfdsqslsvclhevaywyls 660  
QY 661 IGAOTDFLSVFEFSYTRKKHVVYEDTLTFPFSGETYFMSNENGLWILGCHNSDFPRRG 720  
DB 661 igatdfslsvfefsyttrkhhvvyledtltfpfsgetyfmsnenglwiltgchnsdfr 720  
QY 721 MTALLKVSQCDKNTGDYVEDSYEDISAVLTKNNAIEPRFSQNSRHPSTROKOFNAT-- 778  
DB 721 mtaillkvsqcdkntgdyvedsyedisavltknnaieprfsqnsrhpstlrqkfna 778  
QY 779 ----- 778  
DB 781 pendiektpwfahrtmpkignvssdlmlrlqsptphglsldiqeakyetfsdds 840  
QY 779 ----- 778  
DB 841 pgaldsnlssemhfrpqllhsgdmvftpesqlqlrlneklgtateltklqfkvsst 900  
QY 779 ----- 778  
DB 901 smnlstclpsnlagrdntcslqpsmpvnydsqldtllfgkxspstsesgplslsee 960  
QY 779 ----- 778  
DB 961 ndskllsajmsqesswgkuvsteagrlfkqkrahgpalldkdnalfkvislilktn 1020  
QY 779 ----- 778  
DB 1021 ktsnsatnrkthldpslllenspsvwnllesdefkvtpllhdmimdnatalr 1080  
QY 779 ----- 778  
DB 1081 nhmsnktctskmemvqgkkesgripdaqnpdmsffkmlflpesarwlrthgknsinsg 1140  
QY 779 ----- 778  
DB 1141 qgpepkqlvsigpekvsvegqnlseknkvvygkgeftkdvglkemvfpssrnlfltnldn 1200  
QY 779 ----- 778  
DB 1201 lhennthngekkigeleketlligenuvlpqllhvtcgnfknllflstrgnveg 1260  
QY 779 ----- 778  
DB 1261 gayayavldfrslnsdnctrrkthahfskgeeenlsglgnqtkqivekyactrrispnt 1320  
QY 779 ----- 778  
DB 1321 sqgnfvgrskralkgfrlpleetelekrlivddtscqskmhkhlprstltqdyneke 1380  
QY 779 ----- 778  
DB 1381 kgaltspjledcltrshslpqausrpljakvssfpisrplyltrlvfgdnshlpaasy 1440  
QY 779 ----- 778  
DB 1441 rkksdyqesshflgskakmnlslaitltemtdqrevsglgsatnsvykkventvlp 1500  
QY 779 ----- 778  
DB 1501 kpdlptkskvellpkvhiyqkalfpetsngspghidlvsgslvgtegalikwneanrp 1560

QY 779 ----- 778  
Db 1561 gkvpflvatesasaktpsk1ldplawdnhygtq1pkewksqekapektafkktl1sl 1620  
QY 779 ----- PVLAKRIQRETRTLOSDOPE 800  
Db 1621 nacesnhalaiainegnqkpelevtwakqgarterlcsqnpv1krltqreltrtltsdqee 1680  
QY 801 IDYDDTTSVEAMKKEDEFDIYDEDENSPRSFOKTRHYFLAAVERLMDYGMSSSPVLRNR 860  
Db 1681 ldyddtlsvemkkedfdldyedengsprsfqktrhyflaaverlwdgmssphvlmr 1740  
QY 861 AQSQSVPOFKVVFQEFDTGSGFTQPLRYGELNEHLGLGPRYIRAEVDNIMVFRNQSR 920  
Db 1741 agsgsvpqkfvfgeftdgsftqplrygelnehlg1lqpyraevdnlmftfngaar 1800  
QY 921 PYSFYSSLSISEEDROGCAEPRKNVKNPNEKTYFMKVOHNAHPKDEPDKAMAVFSV 980  
Db 1801 pysfyslslsyeedqrgaepkntkvpnektlytwkvqhmapkdecdkawayfsdv 1860  
QY 981 DLEKDVHSGILGPLVCHTNTLNPAGHQVTVQEFALFETIDETKSMYFTENMRNCRA 1040  
Db 1861 dlekdvhsg1lqpllvchntlnpnhgvtvqefalfftlfdektswyftenmerncra 1920  
QY 1041 PCNTIOMEDPTFKENYRFHAINGYIMDTLPGLVMAODRIRWYLSMGSNENHSHFSGH 1100  
Db 1921 pcn1qmedptfkennyrfhalingymdtlpglvmaodqrlrwy1lsmgsnenhshfsgh 1980  
QY 1101 VFYTRKKEEYKMALYNLYPGVFETVEMLPKSKAGIWRVECLIGEHHAAGSTFLVYSKNC 1160  
Db 1981 vflytrkkeeykmalynlypgvfetvemlpksg1wrvecl1gehlhagmstfllyvysnkc 2040  
QY 1161 QTPUGMASGIRDRFOITASGOYGQWAPKLARLHYSGSINAMSTKPPFWIKYDLAPMT 1220  
Db 2041 qtpugmasgh1rdfrdfigtsggygqwapklarlhygsinawstkepfswikyd1lapmt 2100  
QY 1221 HGKIQGAKQKFFSLYISQFTIMWSLDGKMWOTYRGNSTGTMVEFGVNDSSG1KHNFN 1280  
Db 2101 hgk1qgarqkfss1lysqftimwsldgkwmqcyrgnstgtlmvffgnvdsq1khnfn 2160  
QY 1281 PPIIARIYRLHPHTHYSTRSTIRMEIMGCDLNSCAMP1MESKAS1SDAQITASSYTNMRA 1340  
Db 2161 ppi1ariyrlhphtystrstlrmeimgcdlnscamp1meskas1sdag1tassyt1mra 2220  
QY 1341 TWSBSKARLHLQGRSNAMRPQVNNPKEWLQVDFOKTMKVTGVTQGVKSLTSMYKEFL 1400  
Db 2221 twbskarlhlqgrsnawrpqvnmpkewldvdfoktmkvgtvtqgvkslltsmvkefl 2280  
QY 1401 ISSSQDGHQWTLFFQNGKVKVFGQNDSTFPVNSLDPPLLTRYLRIHPQSVHQA1LRM 1460  
Db 2281 isssqdghqwtlffqngkvkvfgqndstfpvnsldppl1trylr1hpqsvwhq1alrm 2340  
QY 1461 EVIGCEAODLV 1471  
Db 2341 ev1gceaqdly 2351

RESULT 15  
AAW13496  
ID AAW13496 standard; Protein: 2351 AA.  
XX AAW13496:  
AC  
XX  
XX  
DT 03-DEC-1997 (first entry)  
DE  
XX Factor VIII:C (Arg1689Lys).  
XX Factor VIII:C analogue; substitution; Factor VIII:C deficiency;  
KW mammal; haemophilia A.  
XX  
OS Homo sapiens.  
XIX

PH Key Location/Qualifiers  
FT Misc-difference 1708  
ET /Label= Arg1689Lys  
PN W09703194-A1.  
XX 30-JAN-1997.  
PD 09-JUL-1996; 96WO-US11441.  
XX 11-JUL-1995; 95US-0001029.  
PR (CHIR ) CHIRON CORP.  
PA  
PI Burke RL, Rosenberg S;  
PI  
XX WPI: 1997-119049/11.  
DR N-PSDB: AAT61548.  
DR  
XX Factor VIII:C analogue with Arg 1689 substituted by Lys - useful for  
PT preventing and treating a Factor VIII:C deficiency, particularly  
PI haemophilia A  
XX  
PS Claim 1; Page -: 46pp; English.  
XX  
CC This sequence represents an active Factor VIII:C polypeptide analogue.  
CC The analogue comprises a native Factor VIII:C polypeptide that is  
CC modified by substitution of the Arg residue at position 1689 with  
CC Lys. The polypeptide analogue can be used to prevent and treat a  
CC Factor VIII:C deficiency in a mammal, esp. haemophilia A. This  
CC sequence is not given in the specification and is based on the  
CC sequence given in Genbank Ref. K01740.  
XX  
SQ Sequence 2351 AA:  
  
Query Match 94.2%; Score 7416; DB 18; Length 2351;  
Best Local Similarity 62.5%; Pred. No. 0;  
Matches 1470; Conservative 1; Mismatches 0; Indels 880; Gaps 1;  
  
QY 1 MOELSTCFEFLCLRFCSATRRRYLGAVELSMQDMSDGLGELPVDAFFPRVPSFPFN 60  
Db 1 mqlstcfeflclrfcsatrrrylgavelswdmsdglgelpvdarfprvpksfpfn 60  
QY 61 TSVVYKKTLEVEFTVHLNIAKPRPPMGLGPTIOAEVDTVYITLKNMASHPSLHAY 120  
Db 61 tsvvykktlveftvhlfniaakprppmglgptioaevdtyv1tlknmashps1hav 120  
QY 121 GVSYWKASGAEYDQTSQREKEDKVPFGSGSHYTWQVLEKENGPMASDPLCLTYSLSH 180  
Db 121 gvsywkasegaeaydqtsqrekedkvfpfgsgshytwvqlkengpmasdp1cltyslsh 180  
QY 181 VDLVLDNSGLIGALVYREGSLAKEKTQTLHKFTLFLPAVDEKSMHSEKNSLMDRD 240  
Db 181 vdlvldnsgl1galvyrsgslakektqtlhkf1flpavdeksmwseknslmdrd 240  
QY 241 AASARAWPKMHTVNGYVNRSLPGLIGCHRSKSYWHVIGMTTPPVHSLFLEGHTFLVRNH 300  
Db 241 aasaraawpkmh1tvngyvnrslpgl1gchrskwywhvigm1tppvhs1fleghtflvrnh 300  
QY 301 ROASLEISPTFFLTAQTLMDLGOFLFCHISRHODGMEAYVYVDSCEPPEQLRMKNNE 360  
Db 301 rgasleisptffltaqtlmdlmgflfch1ssrhodgmeayvrvdscpeepqlrmkne 360  
QY 361 EAEDYDDDLTSEMDYVAFDDDNSSPFOISVAKKHPRKTVHVI1AAEEEDMDVAPLVLA 420  
Db 361 eaedydddltsemdvyafdddnsspf0isvakkhprk1tvhvi1aaeeedmdv1aplvla 420  
QY 421 PDDRYSKQOYLNNQPORIGRKYYKKVRFMAVYDDEFKTRFAAIOHSGILGPLLGEVGDTL 480  
Db 421 pddr1yskqo1l1nnqpor1g1rkyykkvrfmav1yddf1ktrf1aaiohsg1l1gpl1gevgd1l 480  
QY 481 LTIERNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDPFLPGEITFKKWTYTVEDGP 540



```

|||||
Db 481 lllfngsrrpniyphglttdvryrrrlpkvghlkdfrlllpeelkkyktvveqgp 540
OY 541 TKSDBRCLTRYSSVFNMERDLASGLIPBLICYKESVDORGNQMSKRNVIIEFSVEDE 600
Db 541 tksdprcltrrysffvmerdlasglilpbllicykesvdqrgnqmsdkrnvllfsvde 600
OY 601 NRSWLTENIORFLPNPAGVOLEDEPEQASNTMHSINGVFPISLOSLCHVAAWYLLS 660
Db 601 nrswlteniorflpnpagvledepeqasnltmsingvtdslqsvclhevawylls 660
OY 661 IGAQTDPLSVFESGYTFKHKWYEDTLTLFPFSGETVFMSEMPGLWLIGCHNSDFRRNG 720
Db 661 igaqtdplsvfessytfkhwyyedtlclfpfsgetvfmsempglwlilgchnsdfrrng 720
OY 721 M7ALLKVSODCKNCDYEDTEDISAYLLSKNNAIEPPRSQNSRHPSTROKOFNAT-- 778
Db 721 m7allkvsscdkntgdyedsedisyllsknaieprsfqnsrhpstrqkfnatll 780
OY 779 ----- 778
Db 781 pendlekcdpwfahrtcumpkignvssdqlmllqspctphglsldlqeaaketfsdds 840
OY 779 ----- 778
Db 841 pgaldsmnlsencthrpqjlnhsqdmvftpesqqlrlneklgttaateklkldfkvest 900
OY 779 ----- 778
Db 901 snnlstlpsdnlaagtdntesslgppsmvnydsqldtcllfgkkespltesqplslsee 960
OY 779 ----- 778
Db 961 ndeklllesglnmageswgknvssstesgrllfkgrahgpalltkdna lfkvssllktn 1020
OY 779 ----- 778
Db 1021 ktsmsatnrkthidgplllienspsvwgnllesdtefkvtprlindhmlmdknatalrl 1080
OY 779 ----- 778
Db 1081 nhmsnkttskrmemvgqkegplpdpdaqnpdmsffkmlflpesarvigrthgknslnsg 1140
OY 779 ----- 778
Db 1141 qgspkqlvsjgpeksvegnflseknkvvgkgeftkdvglkemvfpssrnlfltnldn 1200
OY 779 ----- 778
Db 1201 lhenmthgekkigeeketlligenvlpqjhtvtgktnfmknlflstrqvegsye 1260
OY 779 ----- 778
Db 1261 gayapvldfrlndstnrctkhtahfakgeeenleglgnqtqivekyactrispnt 1320
OY 779 ----- 778
Db 1321 sqgnfvtrskralqkfrlpleetelekrliivdctstqwmknhltpsclltqidyneke 1380
OY 779 ----- 778
Db 1381 kgaltspqslsdccltrshsldpqnrsplakvssfsplrplylctrvlfgdmsshpaasy 1440
OY 779 ----- 778
Db 1441 rkkdsyvgvesshllqgakknnlsajalllemtgddqrevgslyatsatnsvtkykventvlp 1500
OY 779 ----- 778
Db 1501 kpdlpktsqkveellpknvhyqkdllfptetsngspghldlvegsllqgtegalakwneanrp 1560
OY 779 ----- 778

```

```

Db 1561 gkvflrvateessaktprkllldplawdnhygrtdpkeeksgsekspektafkdkdttllsl 1620
OY 779 -----PVLKRNQOREITRTTLOSDBE 800
Db 1621 nacesnhaalaenegnkpellevtwakggrterlcsqnpvpvlkrrqreltrrltlgsdgee 1680
OY 801 IDYDITISWEMKKEDFDIYDEDEENOSPRSFOKTRHYFIAAVERLMDVGMSSSPVLRNR 860
Db 1681 ldydttiswemkkedfdlydedengspksfgkcltnhyflaaverlwdygmssspvlrnr 1740
OY 861 AOSGSPQFKKVVFOEFTDGSFTQPLRYRGLNEHLGLLGPYIRAEVEDNIMVTFRNQASR 920
Db 1741 aqsgspqfkvvfgefdgstftqplryrgelnehglllgpylraevdnimvtfngasr 1800
OY 921 PYSFYSLSIYEDDROGAEPKRNKVPKNETKTYTWKVOHMAPKDEFDCKAMAYFSDV 980
Db 1801 pysfyslsiyeeddrgaepkntfvknetktyfwkvghmapkdefdckawayfsdv 1860
OY 981 DLEKDVHSGLIIGPLLVCHTNTLNPAGROVTVQEFALPFTIPDETCKSWYFTENMERNCRA 1040
Db 1861 dlekvhnsglllgpllvchntnlnpagnrgvtvqefalffltldektswyfteamerncra 1920
OY 1041 PCNIOMEDPTEKENYRFAINCYIMDTLPGLVMAODORIRMYLLLSGMSNENIHSIHFSCH 1100
Db 1921 pcniomedptfkenyrfaingyimdtlpglvmaqdqrlrwyllsmgsnenihsihfsg 1980
OY 1101 VFTVRRKEEYKMAALYNLPGVEYVEMLPKAGIMRVECLIGEHJHAGNSTLFLVYSNKC 1160
Db 1981 vftvrkkeeykmalynlpgvetyvemlpkagimrvecllgejhagnstlflvysnkc 2040
OY 1161 QTPGLMASGHTRPDQITRSGOVGOMAPKLARLHYSGSINAMSTKBPFSIKYDILLAPMI 1220
Db 2041 qtpglmasghtrpdqitrtsgovgwapklarlhysgslawstkepfswikvdllapmll 2100
OY 1221 HGKTOGAROKFSSLYISOFLIMYSLDGKKMOTYRGNSTGTLMVFFGNVDSGIRHNITFN 1280
Db 2101 hgktggarokfsslyisqflimysldgkkwgtlyrgnstgtlmvffgnvdsqikhnlfn 2160
OY 1281 PPIIARTIRLPHPTHYSINSTLPMELMGCDLNSCSMPLGMEKSAISDAQITTASSTYTNMFA 1340
Db 2161 ppliartirrlphpthysinstlrmelmgcdlncsmplgmeskaisdaqitassystnmfa 2220
OY 1341 TWSPSKARLHLOGRGNAMRPOVNNPKEMLOYDFOKMTKVTGTQGVKSILTLMSVYKEFL 1400
Db 2221 twspskarlhlogrsgnamrpoynnpxewlqvdfqctmkvtgvtctqgvkslltsmvykefl 2280
OY 1401 ISSQDGHQWTLFFONGKVKVFGQNSDFTTPVYNSLDPPLRLRIRIHQSWVWQIALRM 1460
Db 2281 issqdgghqwtlffngkvkvfvqngdsfttpvvnslldpplrlrlyrlhpsvvhqjalrm 2340
OY 1461 EYLGCCEADLY 1471
Db 2341 evlgeceadly 2351

```

Search completed: November 17, 2001, 13:24:24  
 Job time: 133 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 17, 2001, 17:25:18 ; Search time 11937.4 Seconds  
(without alignments)  
6290.589 Million cell updates/sec

Title: US-09-689-430-1  
Perfect score: 7944

Sequence: 1 tggccaccctctctctgcgc.....agggggggggggggggggt 7944

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 10228115 segs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*  
12: gb\_est12:\*  
13: gb\_est13:\*  
14: gb\_est14:\*  
15: gb\_est15:\*  
16: gb\_est16:\*  
17: gb\_est17:\*  
18: gb\_est18:\*  
19: gb\_est19:\*  
20: gb\_est20:\*  
21: gb\_est21:\*  
22: gb\_est22:\*  
23: gb\_est23:\*  
24: gb\_est24:\*  
25: gb\_est25:\*  
26: gb\_est26:\*  
27: gb\_est27:\*  
28: gb\_est28:\*  
29: gb\_est29:\*  
30: gb\_est30:\*  
31: gb\_est31:\*  
32: gb\_est32:\*  
33: gb\_est33:\*  
34: gb\_est34:\*  
35: gb\_est35:\*  
36: gb\_est36:\*  
37: gb\_est37:\*  
38: gb\_est38:\*  
39: gb\_est39:\*  
40: gb\_est40:\*  
41: gb\_est41:\*  
42: gb\_est42:\*  
43: gb\_est43:\*  
44: em\_esthum10:\*  
45: em\_esthum11:\*  
46: em\_esthum12:\*  
47: em\_esthum13:\*  
48: em\_esthum14:\*  
49: em\_esthum15:\*  
50: em\_esthum16:\*  
51: em\_esthum17:\*  
52: em\_esthum18:\*  
53: em\_esthum19:\*  
54: em\_esthum20:\*  
55: em\_esthum21:\*  
56: em\_esthum22:\*  
57: em\_esthum23:\*  
58: em\_esthum24:\*  
59: em\_esthum25:\*  
60: em\_esthum26:\*  
61: em\_esthum27:\*  
62: em\_esthum28:\*  
63: em\_esthum29:\*  
64: em\_esthum30:\*  
65: em\_esthum31:\*  
66: em\_esthum32:\*  
67: em\_esthum33:\*  
68: em\_esthum34:\*  
69: em\_esthum35:\*  
70: em\_esthum36:\*  
71: em\_esthum37:\*  
72: em\_esthum38:\*  
73: em\_esthum39:\*  
74: em\_esthum40:\*  
75: em\_esthum41:\*  
76: em\_esthum42:\*  
77: em\_esthum43:\*  
78: em\_esthum44:\*  
79: em\_esthum45:\*  
80: em\_esthum46:\*  
81: em\_esthum47:\*  
82: em\_esthum48:\*  
83: em\_esthum49:\*  
84: em\_esthum50:\*  
85: em\_esthum51:\*  
86: em\_esthum52:\*  
87: em\_esthum53:\*  
88: em\_esthum54:\*  
89: em\_esthum55:\*  
90: em\_esthum56:\*  
91: em\_esthum57:\*  
92: em\_esthum58:\*  
93: em\_esthum59:\*  
94: em\_esthum60:\*  
95: em\_esthum61:\*  
96: em\_esthum62:\*  
97: em\_esthum63:\*  
98: em\_esthum64:\*  
99: em\_esthum65:\*  
100: em\_esthum66:\*  
101: em\_esthum67:\*  
102: em\_esthum68:\*  
103: em\_esthum69:\*  
104: em\_esthum70:\*  
105: em\_esthum71:\*  
106: em\_esthum72:\*  
107: em\_esthum73:\*  
108: em\_esthum74:\*  
109: em\_esthum75:\*  
110: em\_esthum76:\*  
111: em\_esthum77:\*  
112: em\_esthum78:\*  
113: em\_esthum79:\*  
114: em\_esthum80:\*  
115: em\_esthum81:\*  
116: em\_esthum82:\*

117: gb\_est48:\*  
118: gb\_est49:\*  
119: gb\_est50:\*  
120: gb\_est51:\*  
121: gb\_est52:\*  
122: gb\_est53:\*  
123: gb\_est54:\*  
124: gb\_est55:\*  
125: gb\_est56:\*  
126: gb\_est57:\*  
127: gb\_est58:\*  
128: gb\_est59:\*  
129: gb\_est60:\*  
130: gb\_est61:\*  
131: gb\_est62:\*  
132: gb\_est63:\*  
133: gb\_est64:\*  
134: gb\_est65:\*  
135: gb\_est66:\*  
136: gb\_est67:\*  
137: gb\_est68:\*  
138: gb\_est69:\*  
139: gb\_est70:\*  
140: gb\_est71:\*  
141: gb\_est72:\*  
142: gb\_est73:\*  
143: gb\_est74:\*  
144: gb\_est75:\*  
145: gb\_est76:\*  
146: gb\_est77:\*  
147: gb\_est78:\*  
148: gb\_est79:\*  
149: gb\_est80:\*  
150: gb\_est81:\*  
151: gb\_est82:\*  
152: gb\_est83:\*  
153: gb\_est84:\*  
154: gb\_est85:\*  
155: gb\_est86:\*  
156: gb\_est87:\*  
157: gb\_est88:\*  
158: gb\_est89:\*  
159: gb\_est90:\*  
160: gb\_est91:\*  
161: gb\_est92:\*  
162: gb\_est93:\*  
163: gb\_est94:\*  
164: gb\_est95:\*  
165: gb\_est96:\*  
166: gb\_est97:\*  
167: gb\_est98:\*  
168: gb\_est99:\*  
169: gb\_est100:\*  
170: gb\_est101:\*  
171: gb\_est102:\*  
172: gb\_est103:\*  
173: gb\_est104:\*  
174: gb\_est105:\*  
175: gb\_est106:\*  
176: gb\_est107:\*  
177: gb\_est108:\*  
178: gb\_est109:\*  
179: gb\_est110:\*  
180: gb\_est111:\*  
181: gb\_est112:\*  
182: gb\_est113:\*  
183: gb\_est114:\*  
184: gb\_est115:\*  
185: gb\_est116:\*  
186: gb\_est117:\*  
187: gb\_est118:\*  
188: gb\_est119:\*  
189: gb\_est120:\*

190: gb\_est121:\*  
191: gb\_est122:\*  
192: gb\_est123:\*  
193: gb\_est124:\*  
194: gb\_est125:\*  
195: gb\_est126:\*  
196: gb\_est127:\*  
197: gb\_est128:\*  
198: gb\_est129:\*  
199: gb\_est130:\*  
200: gb\_est131:\*  
201: gb\_est132:\*  
202: gb\_est133:\*  
203: gb\_est134:\*  
204: gb\_est135:\*  
205: gb\_est136:\*  
206: gb\_est137:\*  
207: gb\_est138:\*  
208: gb\_est139:\*  
209: gb\_est140:\*  
210: gb\_est141:\*  
211: gb\_est142:\*  
212: gb\_est143:\*  
213: gb\_est144:\*  
214: gb\_est145:\*  
215: gb\_est146:\*  
216: gb\_est147:\*  
217: gb\_est148:\*  
218: gb\_est149:\*  
219: gb\_est150:\*  
220: gb\_est151:\*  
221: gb\_est152:\*  
222: gb\_est153:\*  
223: gb\_est154:\*  
224: gb\_est155:\*  
225: gb\_est156:\*  
226: gb\_est157:\*  
227: gb\_est158:\*  
228: gb\_est159:\*  
229: gb\_est160:\*  
230: gb\_est161:\*  
231: gb\_est162:\*  
232: gb\_est163:\*  
233: gb\_est164:\*  
234: gb\_est165:\*  
235: gb\_est166:\*  
236: gb\_est167:\*  
237: gb\_est168:\*  
238: gb\_est169:\*  
239: gb\_est170:\*  
240: gb\_est171:\*  
241: gb\_est172:\*  
242: gb\_est173:\*  
243: gb\_est174:\*  
244: gb\_est175:\*  
245: gb\_est176:\*  
246: gb\_est177:\*  
247: gb\_est178:\*  
248: gb\_est179:\*  
249: gb\_est180:\*  
250: gb\_est181:\*  
251: gb\_est182:\*  
252: gb\_est183:\*  
253: gb\_est184:\*  
254: gb\_est185:\*  
255: gb\_est186:\*  
256: gb\_est187:\*  
257: gb\_est188:\*  
258: gb\_est189:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.



QY 5920 gaagatcctttagatcttcttctacgaggctcgaacgtcagtggaacgaacacacgtttaa 5979  
|||||  
Db 601 GAAGATCCTTTGATCTTTTCTACGGGCTGACCGCTCAGTGAAGAAACTCAGCTTAA 660  
QY 5980 gggatttggcatgatataatcaaaaagatcttcacctaagatccctttaataaaaa 6039  
|||||  
Db 661 GGGATTTTGGTCATGAGATATATCAAAAAGATCTTCACTATGATCTTTTAAATTAATAA 720  
QY 6040 tgaagctttaatcaactaagaatataatgaatgaacttgctgacagttaccatgac 6099  
|||||  
Db 721 TGAAGTTTAAATCAATCTAAAGATATATACAGTAACCTTGCTGACAGTTTACCATGTC 780  
QY 6100 ttaacagtaggagcaacctatctcagcagatctgtctattctgttcacatagttgctga 6159  
|||||  
Db 781 TTATCAGTAGGACCACTATCTCAGCATCTGCTATTTGCTCATCATCATGTTGCTTGA 840  
QY 6160 ctcccgctgctgttagataactacgaacagggaggttaccatcttgcccgagtgctga 6219  
|||||  
Db 841 CTCCTCCGCTGTGATGATTAATACGATACGGAGGGGCTTACATCTGCCCCCAGTGTGCA 900  
QY 6220 atgataccgagagaccacgctcacggtccagattatcagcaataaacacgacgacc 6279  
|||||  
Db 901 ATGATTAACGGGAGAACCCACGCTCACCGCTCAGATTTATACGAATTAACACGACGCC 960  
QY 6280 ggaaggccgagcgagcaagtgctcctgcaacttaccgctccatccagctcataat 6339  
|||||  
Db 961 GGAA-GGCCGAGCGCAGAAAGTGCTGCAACTTATCCGCTCCATCCAGCTATTTAAT 1019  
QY 6340 tgttgcgcggaagactagatgaagtgtcgccagttatagttgcgcaac 6390  
|||||  
Db 1020 TGTTCGCGGAGAACTAGATGATGTTGCGCACTTAATGTTTGGCGCAAC 1070

RESULT 2

AJ281480  
LOCUS AJ281480 1004 bp mRNA EST 30-JUN-2000  
DEFINITION 4A3A-P4G8-F Anopheles gambiae immune competent 4A3A Anopheles  
gambiae cDNA clone 4A3A-P4G8, mRNA sequence.  
ACCESSION AJ281480  
VERSION AJ281480.1 GI:6929360  
KEYWORDS EST.  
SOURCE African malaria mosquito.  
ORGANISM Anopheles gambiae.  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoides  
; Anopheles.  
1 (bases 1 to 1004)  
Dimopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C.,  
Donohue, M., Schultz, D., Benes, V., Bork, P., Ansoirge, W., Soares, M.B.  
and Kafatos, F.C.  
Anopheles gambiae pilot gene discovery project: identification of  
mosquito innate immunity genes from expressed sequence tags  
generated from immune-competent cell lines  
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)  
20300950

JOURNAL CONTACT: Dimopoulos G  
MEDLINE FORIS C. Kafatos Laboratory  
COMMENT European Molecular Biology Laboratory  
Meyerhofstrasse 1, 69117 Heidelberg, Germany.  
LOCATION/Qualifiers  
1.1004  
/organism="Anopheles gambiae"  
/strain="4A r/r"  
/db\_xref="taxon:7165"  
/clone="4A3A-P4G8"  
/clone\_lib="Anopheles gambiae immune competent 4A3A"  
/cell\_line="Immune competent 4A3A"  
/lab\_host="E. coli DH10B"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker. Site\_1: EcoRI; Site\_2: NotI; sequenced from  
forward priming site which reads from the 3' end of the  
cDNA. The 4A3A is a directionally cloned and normalized

BASE COUNT 252 a 265 c 244 g 244 t 2 others  
ORIGIN  
Query Match 11.6%; Score 920.8; DB 104; Length 1004;  
Best Local Similarity 98.8%; Pred. No. 1.7e-254;  
Matches 999; Conservative 1; Mismatches 4; Indels 7; Gaps 7;  
cDNA library that was constructed from the 4A3A cell line  
oligo-T primed cDNA according to: Bonaldi, Lennon & Soares  
(1996) : Normalization and Subtraction: Two Approaches To  
Facilitate Gene Discovery, Genome Research 6, 791-806."

QY 5291 aaaagccaggaacgcgttaaaaagcgcgctgtcgtgcttcttcataagctccgccccc 5350  
Db 1 AAAAGGCCAGAACCGTAAAAAAGCGCGCTGTGCGGTTTTCATAGCGCTCCGCCCCC 60  
QY 5351 ctgacgagcatcacaaaatcgacgtcgaagtcgaagttgaggaacccgacagagctat 5410  
Db 61 CTGACGACATCACAAAATCGACGCTCAAGTCAGAGTGAGCGAACCAGACGACTAT 120  
QY 5411 aaagataccagcgcttcccccctggaagctccctcgtgagctcctctgttcgaacctgc 5470  
Db 121 AAAGATACACAGCGCTTCCCGCTGGAAGCTCCCTGAGCGCTCTCGTTCCGACCTCG 180  
QY 5471 cgcttaccggaataccgtgtccgccttctcccttcgggaagcgtagcgcttctcaatgct 5530  
Db 181 CGCTTACCGGATATACCTGTCCGCTCTTCTCCCTTCGGGAACGTCGCGCTTTCATAGCT 240  
QY 5531 cagcgttaggtatctcgaagtcagtgagtgctgtcctcaagctgagtgtagcg 5590  
Db 241 CACGCTGAGGATATCTCAAGTGGGTGAGGCTGCTCCCAAGCTGGGCTGTTTCACG 300  
QY 5591 aacccccgcttcaagcccgacgcgctgcctatccgtaactatcgtctgaattccaac 5650  
Db 301 AACCCCGCTTACCCCGACCGACCGCTGCGCTTATCCGCTATTCGCTGATTCACAC 360  
QY 5651 cgttaagaacagacttaccgcaactgacgacgacactgtaacagattagcagcg 5710  
Db 361 CGGTAAACACACAGCTATTCGCCACTGCGCACAGCACGACACTGTAAACAGATTAGCAGGCA 420  
QY 5711 ggtatgtaggggtgtctacagagttcttgaagtggtgagtcctaactcggtaacagaa 5770  
Db 421 GGTATGAGGGGCTGACACAGATTTTGAAGTGGTGGCTTAACCTAGGCTACACTAGAA 480  
QY 5771 ggaagatttggatctgctgcctcgtcgtgaagcaggttacccttcggaagaaaggttgta 5830  
Db 481 GAACAGATTTTGTATCTGCGCTGCTGTAAGCCAGTTACTTTCGGAAGAAAGTTGGTA 540  
QY 5831 gctctgacccggaacaaacacacgcgctgtgtagcggtgttcttctgttcgaacgac 5890  
Db 541 GCTCTGATCCGGCAAAACCAACCGCTGTAGCGGTGTTTGTTCGCAACGAC 600  
QY 5891 agattaagcagaaaagaaagatcagaagaatccttgccttcttcagaggtctg 5950  
Db 601 AGATTACGCGCAAAAAGATCTCAAGAGTCTTTTGATCTTTTCTACGGGGTCTG 660  
QY 5951 acgctcagtggaagaaacacacacacgcttgaaggatttgcacatgagatlaacaaaagga 6010  
Db 661 ACGCTCAGTGGAGAAAGAACCTCAAGTTAAGGATTTTGGTCATGAGATTATCAAAAAGGA 720  
QY 6011 tcttcacctagatcctttaaataaatgaagttttaaacaactcaagtatatg 6070  
Db 721 TCTTCACCTGATCTTTTAAATTAATAAGATTAAATCAATCAATGATATATG 780  
QY 6071 agtaaatgtgtcgaagattacaaatgcttaactagtgaggaacctatcagagatc 6130  
Db 781 AGTAACTTGGCTGACAGTATCAATGCTTAATCACTGAGCA-CTATCTCAGCATCT 839  
QY 6131 gttcattcttcaatcagatgtgctgactcccgctgctgtagataactacagatacgg 6190  
Db 840 GTCTATTTCGTTCACTCATGATGCTGACTCCCGCTGCTGAT-ACTACGATACGGG 898  
QY 6191 agggttaccatctgcccagtgctgcaatgataccgcgaagaccacgctcaccgctc 6250

```

|||||
Db 899 AGGCGTACCATYTG-CCAGTGCTGCAATGATACCGCAGAG-CCAGCTCA-CGCGTC 955
QY 6251 cagattatagcaataaaccagccggaagggcgagcgagcgagcg 6301
Db 956 CAGATT-TGAGCAATTAACCGACCGCAGGAA-GGCGGAGCGCAGAGTG 1004

RESULT 3
AL042026/c 841 bp mRNA EST 29-FEB-2000
LOCUS AL042026
DEFINITION DKFZp434E11.1 434 (synonym: htes3) Homo sapiens cDNA clone
AL042026
ACCESSION AL042026
VERSION AL042026.1 GI:5421372
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 841)
AUTHORS Poustka,A., Klein,M., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE EST (Poustka, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Poustka A.J.
Department Lehnach
Max-Planck-Institute for Molecular Genetics
Imestrasse 73, 14195 Berlin, Germany
Tel: +49-30-84131623
Fax: +49-30-84131128
Email: poustka@mpimg-berlin-dahlem.mpg.de
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No sl sequence available.
This clone (DKFZp434E11) is available at the RZPD in Berlin
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1. 841
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp434E11"
/clone_lib="434 (synonym: htes3)"
/issue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

BASE COUNT 226 a 192 c 202 g 221 t

ORIGIN
Query Match 10.6%; Score 841; DB 105; Length 841;
Best Local Similarity 100.0%; Pred.No. 1.9e-231;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6204 tggcccgatgtctgaatgatalaccgagagccacgctccagctccagattatcagc 6263
Db 841 TGCGCCAGTGTGCAATGATACCGGAGACCGACGCTCACGGCTCCAGATTATCAGC 782
QY 6264 aataaacccagcagcggaaggcgagcgagaaagtgtctcgtcaacttaccgcctc 6323
Db 781 AATTAACCAACCCAGCGGAGGCGCCAGCCAGCAAGTGTGCTCTTAATCCCGCTC 722
QY 6324 catcagctcataatgtgtgctcggaagagtagaagtaagtaagtcgccagttatagttt 6383
Db 721 CATCCAGTCTAATTAATGTGTGCGGGAAGCTAAGTAAGTATCCCACTTAATAGTTT 662
QY 6384 ggcgaacgttgtctcatgtctacagcagcagtcgtgtcagcgtcgttgtgtatgac 6443
|||||

```

```

Db 661 GCCCAAGTGTGTCATCTACAGCATGCTGTGTCAAGCGCTGCTGTGTATGCG 602
QY 6444 ttcatcagctccggttcccaacgataaagcgagtaagtagtcccatgtgtgcaa 6503
Db 601 TTCAATTCAGCTCCGGTCCCAACGATCAAGCGAGTTACATGATCCCGCATGTGTGCAA 542
QY 6504 aaaaagcgttagctccttcgctcccgatcgtgtgtcaagaagtaagtgtgacgagttc 6563
Db 541 AAAAGCGGTAGCTCCTTCGCTCTCGATCGTGTTCAGAGTAAGTGGCCGACGTGT 482
QY 6564 atcaactatggtatgagcaagcaactgataatctctactgltcagtcacglaagatg 6623
Db 481 ATCAGTCATGTTATGGCAGCAGCTGATATTTCTTACTAGTCATGCGATCCGTAAGATG 422
QY 6624 ctttctgtgactgtgtgagtaactcaacaaagtaactctgagaatagtgtatgcggacc 6683
Db 421 CTTTCTGTGACGTGTGAGTACATCAACCAAGTCATTTGTGAATAGATGTATGCGGAC 362
QY 6684 gagtgcctctgcccggcggtcaataagggatlaacggcgacacatagcagaactttaa 6743
Db 361 GAGTTGCTCTTGCCCGCGCTCAATACGGGATTAATACCGCCACATACGAACTTTAAA 302
QY 6744 agtgcacatcatggaanaacgctcttcgaggcgaaaactcgaagatcttaccgctgt 6803
Db 301 AGTGTCATCATTTGGAANAACGTTCTTGCGGGCGAANAACGTCAAGATCTTACCGTGT 242
QY 6804 gagatccagttcgatgttaaccacactcgtgcacccaactgattcttcaagatcttactt 6863
Db 241 GAGATCCAGTTGATGATGACCACTGTCGCCACCACTGATCTTCAGCATCTTTACTTT 182
QY 6864 caccagcgttctcgggtgagcaaaaacaggaagcaaatgcccgaanaaagggaatag 6923
Db 181 CACCAGCGTTTCTGGGTGAGCAAAAACAGCAAGCAAAAATCCGCCAATAAGGAAATAG 122
QY 6924 ggcgaacaggaatgtgtaactacactctctcttcttcaatlatattgaagcatla 6983
Db 121 GCGGACACGGAAATGTGATATCATATCTCTTCCCTTCAATATATTTGAAGCATTTA 62
QY 6984 tcagggtattgtctcatgagcgagatataatttgatgtatttagaanaataacaaat 7043
Db 61 TCAGGGTATTATGTCATGAGCGGATACATATTGATTTAGAAAAATAACAAAT 2
QY 7044 a 7044
Db 1 A 1

RESULT 4
AJ281449
LOCUS AJ281449
DEFINITION 4A3A-P4D5-F Anopheles gambiae immune competent 4A3A Anopheles
gambiae cDNA clone 4A3A-P4D5, mRNA sequence.
ACCESSION AJ281449
VERSION AJ281449.1 GI:6929329
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae
; Anopheles.
1 (bases 1 to 800)
REFERENCE Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C.,
Donohue,M., Schultz,J., Benes,V., Bork,P., Ansoerge,W., Soares,M.B.
and Kafatos,F.C.
Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
JOURNAL MEDLINE
COMMENT Contact: Dimopoulos G
Fotis C. Kafatos Laboratory
European Molecular Biology Laboratory
Meyerhofstrasse 1, 69117 Heidelberg, Germany.

```





```

Db      559 GCGAAGCCGACAGACTATAAGATACCAAGCGTTCCCTCGAAGCTCCCTGTCGC 618
        5451 ctctccgttccgaacctgcgcgttaccgatactctgcgccttctcccttcgggaag 5510
        619 CTCGCTGTTCCGACCGCTGCGCTTACCGATACCTTCGCCCTTCTCCCTTGCGGAG 678
        5511 cgtgagcttctcaatgcaacgctgtagtattcagttcgtgtagtgcgtc 5570
        679 CGTGGCCCTTTCATATAGCTCAGCGCTAGTATCTAGTTCCGTGTAAGTCGTTCCCTC 738
        5571 caagctgagctgtagtgcgaacccccgcgttcagcccgacgcgtcgccttalcggtaa 5630
        739 CAAGCTGGGCTGTGTGACCAACCCCGTTACGCCGCGCTGCGCCTTATCCGGTAA 798
        5631 ctatcgcttctgagtcgaaccccggttaagaacagacttatgcacatggtgcgaacgcaactg 5690
        799 CTATCGCTTGTGAGTCAACCCCGTAAACAGACTTATCCCATGTGACACAGCCACTGG 858
        5691 taacagattagcagaagcagaagtatgtaggcgtgtctacagagttcttgaagtgtggcc 5750
        859 TAACAGGATTAGCAGAGCGAGGATGTAGCGGCTACAGAGTTCTTGAAGTGTGGCC 918
        5751 taactacgctactacagaagaagcaagtattggtatctgcgtctgtgaagccaagtac 5810
        919 TAACTAGGCTACACTAGAGAGACAGATTTGGTATCTGCGCTCTGCTGAAGCCAGTTAC 978
        5811 ctccggaagaaaggttggtagctcttctgattccggaacaaacacacgcgtgtgtgagtg 5870
        979 CTTCGGAAAAAGATTGTGAGCTCTTGTATCCGCCAACAACAA-CAACGCTGTGATGCGTTT 1037
        5871 ttttttcttgcgaagcagattacgcgc 5901
        Db      1038 TTTTGTGTTTGC-AGCAGCAGATTACGCCG 1067

RESULT 6
LOCUS   AL044364          954 bp      mRNA      EST      29-FEB-2000
DEFINITION DKEFZ434C172_s1 434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION AL044364
VERSION   AL044364.1 GI:5432586.
KEYWORDS  EST.
SOURCE    human.
ORGANISM Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE Ansoerge,W., Benes,V., Krieger,S., Mewes,H.W., Gassenhuber,J. and
AUTHORS   Wiemann,S.
           EST (Ansoerge, Benes, et al.)
TITLE     Unpublished (1999)
JOURNAL   Contact: Ansoerge W
COMMENT   MIPS
           Am Kioferplatz 18a D-82152 Martinsried, Germany
           This is the 3' sequence of the clone insert
           Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
           Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
           Heidelberg/Germany) within the cDNA sequencing consortium of the
           German Genome Project.
           No r1 sequence available.
           This clone (DKEFZ434C172) is available at the RZPD in Berlin
           Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
           Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
           Location/Qualifiers
             1..954
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone="DKEFZ434C172"
             /clone_lib="434 (synonym: htes3)"
             /tissue_type="testis"
             /dev_stage="adult"
```

```

BASE COUNT      214 a      273 c      253 g      214 t
ORIGIN
Query Match      9.7%; Score 769.2; DB 105; Length 954;
Best Local Similarity 99.6%; Pred. No. 1.2e-210;
Matches 771; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
/!ab host="DH10B"
/!ote="Vector: pSport1; Site.1: NotI; Site.2: SalI"
Db      5095 ctgcattaaagaaatcgcgcgaacgcgcgcgggagagcggttgcgtatctggcgctctcc 5154
        181 CTGATTAATGAAATCGGCCCAACGGCGGGAGAGCGCGTTTGGGATTTGGCGCTCTTCC 240
        5155 gcttctcgctcaactgactgcgtcgtgcgttgcgttgcgttcgctgcgcgcgcgcgtatcaagt 5214
        241 GCTTCCCTCCGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
        5215 cactcaaaagcgcgtlaaagcgttataccagaatccagggatatacgaagaaagaaacatg 5274
        301 CACTCAAGGCGGATATACGATTATCCACAGAAATCAGGGGATTAACGAGAAAGAACATG 360
        5275 tgaacaaaaggcagcaaaagccaggaacgcgtlaaaaggccgcgctgtgcgtcttctc 5334
        361 TGAGCAAAAGCGCCAGCAAAAGCCAGGAACCGTAAAAAGCCCGGTTGCTGGCGTTTTC 420
        5335 catagctccgcgcgcctcgaagcatcacaaaatgcagctcaagtcaagagtgcga 5394
        421 CATAGGCTCCGCCCCCTCGACGACATCAACAAAATCGACGTCATGAGAGTGGCGGA 480
        5395 aaccgcgaagactataaagaataccaagggttcccccctgaaagctccctcgtgcgtct 5454
        481 AACCCGACAGGACTATTAAGATACAGGCGTTTCCCTGGAAGCTCCCTGAGCGGCTCT 540
        5455 cctgtccgaacctgcgcgttaccgatacctgtccgccttctcccttcgggaagcgtg 5514
        541 CTTGTTCCGACCGCTGCGGCTTACCGGATACCTGTCCGCTTCTCCTTGCGGAAGCGTG 600
        5515 ggcgttctcaatgctcaagcgtgtagtattcgaatcagttcgtgtagtgctgcctccaag 5574
        601 GCGCTTCTCATGAGCTCACCGCTGAGTATCTCAGTTCCGCTGATGAGTCCCTCCCAAG 660
        5575 ctgggcctgtgtgcaagcaacccccgttcagcccgacgcgcgcgttatccgttaactat 5634
        661 CTGGGCTGTGTGACAGAACCCCGCTTCAAGCCGACCGCTGCGCTTATCGGTAACATAT 720
        5635 cgctctgagtcgaacccgcgtgaagaacagactatcgcacactggaagcgaagcactgttaac 5694
        721 CGCTTGAAGTCCAAACCGGTAACACAGACTTATTCGCCACTGGCAGAGGCACTGTGAAC 780
        5695 aggaattagcagaagcgaagtatgtagcgtgtctacagaagttcttgaagtgtgacctaac 5754
        781 AGGATTAGCAGACGAGGTATGTAGGGGGGTACAGAGTTCTTGAAGTGTGGCTGAAC 840
        5755 taagcctaacctagaagagcaagfatattggtatctgcgcgttcgttcgaagcagttacttc 5814
        841 TACGCTACACTTGAAGAACAGATTTTGGTATCTGCGCTTCTGCTGAAGCCAGATTACCTTC 900
        5815 ggaataaagatgtgtgactctctgattccgcggaacaaacacgcgtgtgtagcgtt 5868
        901 GGAATAAAGATTGTGATGCTTGTATCCGGCAACAAACACCGCTGTGACCGGT 954

RESULT 7
LOCUS   AU081124          1089 bp      mRNA      EST      15-NOV-1999
DEFINITION AU081124 Oncorhynchus mykiss Kidney infected by infectious
ACCESSION AU081124
VERSION   AU081124.1 GI:6431472
KEYWORDS  EST.
SOURCE    rainbow trout.
```

|                       |   |            |           |           |          |          |          |          |              |      |
|-----------------------|---|------------|-----------|-----------|----------|----------|----------|----------|--------------|------|
| ORGANISM              | Oncorhynchus mykiss<br>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus. 1 (bases 1 to 1089)<br>Sakai,M. and Kono,T.<br>The EST analysis of kidney and gill cells rainbow trout, Oncorhynchus mykiss, infected with infectious hematopoietic necrosis virus<br>Unpublished (1999)<br>Contact: Masahiro Sakai<br>Faculty of Agriculture<br>Miyazaki University<br>1-1 nishi kakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan<br>Email: a0b20b@cc.miyazaki-u.ac.jp.<br>Location/Qualifiers<br>1. 1089 |            |           |           |          |          |          |          |              |      |
| JOURNAL               |   |            |           |           |          |          |          |          |              |      |
| COMMENT               |   |            |           |           |          |          |          |          |              |      |
| FEATURES              | source<br>1. 1089<br>/organism="Oncorhynchus mykiss"<br>/db_xref="taxon:8022"<br>/clone="KG12"<br>/clone_1lb="Oncorhynchus mykiss kidney infected by infectious hematopoietic necrosis virus"<br>/tissue_type="Kidney infected by infectious hematopoietic necrosis virus"  |            |           |           |          |          |          |          |              |      |
| BASE COUNT            | 245 a 297 c 295 g 251 t 1 others  |            |           |           |          |          |          |          |              |      |
| ORIGIN                |   |            |           |           |          |          |          |          |              |      |
| Query Match           | 9.3%; Score 741.2; DB 107; Length 1089;   |            |           |           |          |          |          |          |              |      |
| Best Local Similarity | 96.5%; Pred. No.1.6e-202;   |            |           |           |          |          |          |          |              |      |
| Matches               | 801; Conservative 0; Mismatches 23; Indels 6; Gaps 4  |            |           |           |          |          |          |          |              |      |
| OY                    | 5091  | ccagctgcat | aatgaatga | atcgccac  | gcgcg    | gggaa    | ggcgttgc | gtatctgg | gcgcctc      | 5150 |
| Db                    | 260   | CCAGCTGCA  | TATGATGA  | TGCGCCAA  | CGCGGGA  | AGGCGG   | GTTCGAT  | TGGCGCTC |              | 319  |
| OY                    | 5151  | ttccgctt   | ctctgcgt  | ctactgcgt | cgctgtgc | gtgcgttc | gttcggtc | gtgcggag | cgtatc       | 5210 |
| Db                    | 320   | TTTCGGCT   | TCTCGCTCA | GTGAC     | TGCGTGG  | CTGCTG   | TTCCGG   | CTCGG    | GACGGGTATC   | 379  |
| OY                    | 5211  | agctcact   | aaagcgg   | taatac    | ggtatcc  | acagata  | tcagggat | caacg    | caagaa       | 5270 |
| Db                    | 380   | AGCTCACT   | CAAAAGCG  | GTATAC    | GGTTAT   | CCACAGAT | CAAGGG   | GATRAC   | CCAGAAAGAA   | 439  |
| OY                    | 5271  | catlgtag   | caaaagc   | icagca    | aaaggc   | caagaa   | cgtla    | aaaagcc  | gctltgc      | 5330 |
| Db                    | 440   | CATGTGAC   | CAAAAAG   | CGCACAA   | AAAGCC   | ACAGAAC  | CGTTAA   | AAAGCG   | CGCTGTGCGCTG | 499  |
| OY                    | 5331  | tttccata   | gagcttcc  | ggccccc   | ctgcag   | agatcat  | caaaaat  | tcagcgt  | ctcaagt      | 5390 |
| Db                    | 500   | TTTTCCAT   | TAGGCTCC  | GGCCCC    | CTTGAC   | GAGATCC  | CAAAAAT  | TCGAGCT  | CAAGTCAAG    | 676  |
| OY                    | 5391  | gcgaac     | cccgaca   | gagata    | laaaga   | taccag   | gcgttcc  | ccctg    | gaagctt      | 5450 |
| Db                    | 560   | GGGAAC     | CCCGAC    | AGACT     | TTAAAG   | ATACCA   | GCGGTTT  | CCCC     | CTGGAA       | 619  |
| OY                    | 5451  | ctctcc     | gttccg    | accctgc   | gcgttacc | gcgata   | ctcgttc  | gcgcctt  | ctccctt      | 5510 |
| Db                    | 620   | CMCTCTGT   | TCGAC     | CCCTCG    | CGCTTAC  | CGGATAC  | GTCCGT   | CGCCTT   | CTCCGGA      | 679  |
| OY                    | 5511  | cgttgc     | gagctt    | ctctca    | tgtc     | tcaagc   | gtgtat   | ctcc     | agttc        | 5570 |
| Db                    | 680   | CGTGGG     | CGTTTCA   | ATGCTCA   | GGCTGT   | AGGTAT   | CTCAGT   | TTCGGT   | TAGGTCTT     | 739  |
| OY                    | 5571  | caagct     | ggagct    | gtgtgc    | acga     | accccg   | cttaag   | cccgag   | ccgtgtgc     | 5630 |
| Db                    | 740   | CAAGCTGG   | CGCTGTGT  | GCACGA    | ACCCCC   | CGTTCAG  | CCCGAC   | CGCGTGG  | CGCCTTAT     | 799  |
| OY                    | 5631  | ctatcgt    | ctctgag   | tcca      | aacccg   | taaga    | caagctat | tcgcac   | ctgtgc       | 5690 |
| Db                    | 800   | CTATCGTC   | TTGTA     | GTCCAA    | CCCGGT   | TAAGAC   | ACAGACT  | TATCC    | CCACTG       | 859  |
| OY                    | 5691  | taacag     | attag     | tcag      | agcag    | agtata   | gtatg    | agcgttgc | tacagagt     | 5749 |

[illegible]

|            |  |   |             |
|------------|--|---|-------------|
| Db         | 72   | ATCCACAGAAATCAGGGGATATACCCAGGAAAGAACATGTGACGAAAAAAGGCCACACAAAAGGC | 131         |
| Qy         | 5298   | caggaaacgcgttaaaaaagcgcgctgtgtgcgtgttttccalbaggtctccgccccctgcagca | 5357        |
| Db         | 132  | CAGGAACCCGTAAAAAAGGCCCGCTTGGCTGGGNNNTTTCATAGAGCTCCGCCCTCGAGACA    | 191         |
| Qy         | 5358   | gcatacaaaaaatctgagcgccaagttaaggtgtggtggaaccccgacaagacataaagata    | 5417        |
| Db         | 192  | GCATACACAAAATTCAGAGCTCAAGTCAGAGAGTGGCAAAACCCGACAGGACTCTTAAAGATA   | 251         |
| Qy         | 5418   | ccagagcttcccccttggaaagctccccctgtgagctctctctgtttccgaacctgcgcttac   | 5477        |
| Db         | 252  | CCAGAGCTTTTCCCTCTGGAAGCTCCTCGTGGCCTCTCTCTGTTCCGACACCTCCGCCCTTAC   | 311         |
| Qy         | 5478   | cgaatacctgtlccgcgctcttctcccttcgggaagcgtgacgtcttctcaaltgctacgctg   | 5537        |
| Db         | 312  | CGATATACCTGTCGCGCTTTCTCCCTTTCGGGAACGTCGGCGCTTTCATAGCTCACGCTG      | 371         |
| Qy         | 5538   | taggtatcctcaagttcgggttgaagtcgtttcgtctccaagctggagctgtgtgacgaacccc  | 5597        |
| Db         | 372  | TAGTATCTCATGTCGGGTAGGCTGCTGCTGCTCCACAGCTGGGCTGTGTGACAGAACCCCC     | 431         |
| Qy         | 5598   | cgttcaaccccgagacgcgtgagcccttaccggtiaactatgctcttgtagtccaaccggttaag | 5657        |
| Db         | 432  | CGTTCACCCCGACCGCGCTGCGCTTATCCGATCACTATCTGCTTGGAGTCCAAACCCCGTAG    | 491         |
| Qy         | 5658   | acaagacttatcgccactgtgcagcagcacactggtlaacaagattagacagcagtatgt      | 5717        |
| Db         | 492  | ACAGACTTATTCGCACTGCGACACACCACTGATACAGATTACAGACGACAGGTATGT         | 551         |
| Qy         | 5718   | agagcgtgtctacagagattcttgaagtgtgtggccttaactaaggctaacataagaagacgt   | 5777        |
| Db         | 552  | AGGGCGGTCTACAGAGAGTCTTGAAGTGGTGGCTACTACGTACGCTACACTAGAAAGAACGT    | 611         |
| Qy         | 5778   | attgtgattctgggcctctgttgaagcaattactctcgaaaaaagaattgtgaagctctg      | 5837        |
| Db         | 612  | ATTGTGATCTGGCGCTCTGCTGAAGCCAGTACTTCCTGGAAAAAGAGTTGGTAGCTTGG       | 671         |
| Qy         | 5838   | atccgcgaacaacaacacacgcctgtatgcg                                   | 5866        |
| Db         | 672  | ATCCGGCAAAACCAACCCACCGCTGTGAGCG                                   | 700         |
| RESULT     | 9  |   |             |
| LOCUS      | AO856202/c   |   |             |
| DEFINITION | nhebd001.F05f CUGI Rice BAC Library (EcoRI) Oryza sativa genomic   | GSS   | 03-NOV-1999 |
| ACCESSION  | AO856202   |   |             |
| VERSION    | AO856202.1   | GI:620659   |             |
| KEYWORDS   | GSS.   |   |             |
| SOURCE     | Oryza sativa.  |   |             |
| ORGANISM   | Oryza sativa.  |   |             |
| REFERENCE  | Enkayota; Vitridiplantae; Streptophyta; Embryophyta; Tracheophyta;   |   |             |
| TITLE      | Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;   |   |             |
| AUTHORS    | Enhratioideae; Oryzeae; Oryza.   |   |             |
| JOURNAL    | 1 (Bases 1 to 757)   |   |             |
| COMMENT    | Wing, R.A. and Dean, R.A.<br>A BAC End Sequencing Framework to Sequence the Rice Genome<br>Unpublished (1998)<br>Contact: Wing RA<br>Clemson University Genomics Institute<br>Clemson University<br>100 Jordan Hall, Clemson, SC 29634, USA<br>Tel: 864 656 7288<br>Fax: 864 656 4293<br>Email: rwing@clemson.edu<br>Seq primer: TAATACGACTCACTATAGGG<br>Class: BAC ends<br>High quality sequence start: 29<br>High quality sequence stop: 232.<br>Location/Qualifiers<br>1..757 |   |             |

/organism="Oryza sativa"  
 /strain="Japonica"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:4530"  
 /clone="abeb0001f05f"  
 /clone\_1fb="CGRI Rice BAC Library (EcoRI)"  
 /tissue\_type="leaf"  
 /lab\_host="E. coli DH10B"  
 /note="Vector: pBACindigo; Site\_1: EcoRI; Site\_2: EcoRI;  
 Rice is the most important food crop in the world. Half of  
 the world population, especially those inhabiting highly  
 populated areas of the humid tropics and subtropics, rely  
 on rice as their primary source of carbohydrate.  
 Monocotyledonous rice is a diploid plant (2n=24) with a  
 haploid genome equivalent of 431 Mbp (Arunagathan and  
 Earle, 1991). The relatively small genome of rice, three  
 times larger than that of Arabidopsis, makes it suitable  
 for genomic studies. In order to facilitate positional  
 cloning, physical mapping and genome sequencing of rice,  
 we have constructed a BAC library from *Oryza sativa*,  
 Nipponbare variety using EcoRI as the cloning enzyme. The  
 library contains 55,296 clones with an average insert size  
 of 121 Kb providing approximately 15 haploid genome  
 equivalents. The deep coverage allows the isolation a  
 particular sequence with a probability of 99.9 %. Three  
 high density filters, each containing 18,432 clones  
 (doubly spotted), represent the whole library for colony  
 screening and can be requested from the Clemson University  
 BAC/EST Resource Center ([www.genome.clemson.edu](http://www.genome.clemson.edu))."  
 180 c 178 g 185 t 10 others

|                           |        |                     |           |             |
|---------------------------|--------|---------------------|-----------|-------------|
| Query Match               | 8.5%;  | Score 674.6;        | DB 234;   | Length 757; |
| Best Local Similarity     | 98.0%; | Pred. No. 2.7e-183; |           |             |
| Matches 702; Conservative | 0;     | Mismatches 12;      | Indels 2; | Gaps 2;     |

|    |      |   |       |
|----|------|---|-------|
| Qy | 5434 | ggaagctccctcgtgcgtctctctgttcgcagccctgcgcttacgcgatacc-tgtccgc      | 54322 |
| Db | 755  | GAAGGCTCTTTCGGCCCTCTCTGTTCNACNCTCGGCTTTACGGAAACNTTTCNCG           | 696   |
| Qy | 5493 | ctctctcc-cctcggagaagctgtagcctctctcaatgctcaacgcgtgtagatctcagtt     | 55511 |
| Db | 695  | CTTTCCTCCNCTTGGGAAAGGAGGGCGCTTTCATATAGCTACAGCTGTAGTATCTCAATT      | 636   |
| Qy | 5552 | cggtgtcagttcgtctcgtctcaagcttggtgctgtgtcagaacccccgltcaagcccgac     | 56111 |
| Db | 635  | CGGCTGTAGGTCGTTCGGCTCCAACTGGGCTGTGTGCACAAACCCCGCTTCAGCCGACAC      | 576   |
| Qy | 5612 | gctgcagcttatccggtgaactatcgtctctgagtcgaacccggttaagaacagctatcgc     | 56711 |
| Db | 575  | GCCTGGCCTTATCCGGTAACTATTCGTCTTGAGTCCACCCGGTAAAGACACACTTATGCG      | 516   |
| Qy | 5672 | cactgtagcagcagcccaactgltacaagaattagaacagcgaagtatgtaaagcgtgtctacag | 57311 |
| Db | 515  | CACGTGCAGACAGCCACTGGTAAACAGGATTATGACAGACGAGTATAGCGCGTGTCTACAG     | 456   |
| Qy | 5732 | agttcttgaaagtgtgagcccaactacagctcactagaacagatattgtctatcgtcgc       | 57911 |
| Db | 455  | AGTTCTTGAAGTGTGTGGCCCTTAACCTACGCGCTACCTAGAAAGACATATTGTGTATCTCGC   | 396   |
| Qy | 5792 | ctctgtctgaagccagttacctctcggaaaaagattgtgtacgtctctgtatccgcggcaaacaa | 58511 |
| Db | 395  | CTCTGCTGAAGCCAGTTACTCTTCGGAAAAAGATTGTGAGCTCTTGATCCGGCAAAACAA      | 336   |
| Qy | 5852 | ccacagcgtgtgtagcgtgtgatttttttgcgttcggaagaagaagattacgcgcgaaaaaag   | 59111 |
| Db | 335  | CCACCGCTGTAGCGGTGTTTTTTTGTTCGCAACGACAGAGATTTCGCGCGCAAAAAAAG       | 276   |
| Qy | 5912 | gattctcaagaagaatcctttgtatcttcttctacaggggtcgtacgctcagtgtgaagaaact  | 59711 |
| Db | 275  | GATCTCAAGAAGATCTCTTGTATCTTTTCTACGGGGGTCTACCTCATGTGGAACGAAACT      | 216   |

| Y                         | 5972  | caacgttaaggagcttcgtgcatgagatatacaaaaaggagcttcaacctgacatctttaa  | 6031 |     |             |
|---------------------------|---|--|------|-----|-------------|
| Db                        | 215   | CACGTTTAAGGGATTTTGGTCATGACATTATCAAAAAGATCTTCACACTTGATCCTTTTAA  | 156  |     |             |
| Y                         | 6032  | altaaanaatgaagctttaataatcaatctaaagatatalatgagtaaaccttgcgcagtt  | 6091 |     |             |
| Db                        | 155   | ATTAAAAATGAGAGTTTAAATCAATCAATCAATATATATATAGTAAACTTGGCTGCACGTT  | 96   |     |             |
| Y                         | 6092  | accatgtctaatcgttgtaggcacctatctaaaggatcgtctatcttgcctatcc        | 6147 |     |             |
| Db                        | 95  | ACCAATGCTTAAATCAAGTGAAGGACCACTATCTCAGCATGCTATTTCTGCTATCC       | 40   |     |             |
| RESULT 10                 |   |  |      |     |             |
|                           | AJ281437  | 703 bp   | mRNA | EST | 30-JUN-2000 |
| LOCUS                     | AJ281437  |  |      |     |             |
| DEFINITION                | 4A3A-P4C3-F Anopheles gambiae immune competent 4A3A Anopheles             |  |      |     |             |
| ACCESSION                 | gambiae cDNA clone 4A3A-P4C3, mRNA sequence.                              |  |      |     |             |
| VERSION                   | AJ281437  |  |      |     |             |
| KEYWORDS                  | AJ281437.1 GI:6929317   |  |      |     |             |
| SOURCE                    | EST.  |  |      |     |             |
| ORGANISM                  | African malaria mosquito.   |  |      |     |             |
|                           | Anopheles gambiae   |  |      |     |             |
|                           | Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;             |  |      |     |             |
|                           | Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae       |  |      |     |             |
|                           | ; Anophelinae.  |  |      |     |             |
| REFERENCE                 | 1 (bases 1 to 703)  |  |      |     |             |
| AUTHORS                   | Dimopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C.,      |  |      |     |             |
|                           | Donohue, M., Schultz, J., Benes, V., Bork, P., Ansoorge, W., Soares, M.B. |  |      |     |             |
|                           | and Katatos, F.C.   |  |      |     |             |
|                           | Anopheles gambiae pilot gene discovery project: identification of         |  |      |     |             |
|                           | mosquito innate immunity genes from expressed sequence tags               |  |      |     |             |
|                           | generated from immune-competent cell lines                                |  |      |     |             |
|                           | Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)                   |  |      |     |             |
| COMMENT                   | Contact: Dimopoulos G   |  |      |     |             |
| MEDLINE                   | Fotis C. Katatos Laboratory   |  |      |     |             |
|                           | European Molecular Biology Laboratory                                     |  |      |     |             |
|                           | Meyerhofstrasse 1, 69117 Heidelberg, Germany.                             |  |      |     |             |
| FEATURES                  | Location/Qualifiers   |  |      |     |             |
| source                    | 1..703  |  |      |     |             |
|                           | /organism="Anopheles gambiae"   |  |      |     |             |
|                           | /strain="4A r/F"  |  |      |     |             |
|                           | /db_xref="taxon:7165"   |  |      |     |             |
|                           | /clone="4A3A-P4C3"  |  |      |     |             |
|                           | /clone_lib="Anopheles gambiae immune competent 4A3A"                      |  |      |     |             |
|                           | /cell_line="immune competent 4A3A"  |  |      |     |             |
|                           | /lab_host="E. coli DH10B"   |  |      |     |             |
|                           | /note="Vector: p7T3D-Pac (Pharmacia) with a modified                      |  |      |     |             |
|                           | polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from                   |  |      |     |             |
|                           | forward priming site which reads from the 3' end of the                   |  |      |     |             |
|                           | cDNA. The 4A3A is a directionally cloned and normalized                   |  |      |     |             |
|                           | cDNA library that was constructed from the 4A3A cell line                 |  |      |     |             |
|                           | OL30-T primed cDNA according to: Bonaldo, Lennon & Soares                 |  |      |     |             |
|                           | (1996) : Normalization and Subtraction: Two approaches To                 |  |      |     |             |
|                           | Facilitate Gene Discovery, Genome Research 6, 791-806."                   |  |      |     |             |
| BASE COUNT                | 159 a 204 c 193 g 147 t   |  |      |     |             |
| ORIGIN                    |   |  |      |     |             |
| Query Match               | 8.3%; Score 662.4; DB 104; Length 703;                                    |  |      |     |             |
| Best Local Similarity     | 98.4%; Pred. NO. 8.8e-180;  |  |      |     |             |
| Matches 669; Conservative | 0; Mismatches 11; Indels 0; Gaps 0;                                       |  |      |     |             |
| Y                         | 5097  | gcatatgatcaatccggcaacgcgcggcggaagcgaggttgcgtatttgaggcgtcttcgcg | 5156 |     |             |
| Db                        | 24  | GGAAATAGCTTCCGGCCACACGCGCGGGAAGAGCGATTGCGATTGGCGCTCTTCCGC      | 83   |     |             |
| Y                         | 5157  | ttcctcgtctcaactgactcgtctgcgttcggttcgtctgcgtcgcgcagcgtatcagctca | 5216 |     |             |
| Db                        | 84  | TTCTCTCGCTACACTGACTCGCTCGCTCGGTCTTGGCGTGGCGGAGCGGTATCAGTCA     | 143  |     |             |
| Y                         | 5217  | ctcaaaagcggtataacggttatccacagaatcaggggataaagcagagaagaacatgty   | 5276 |     |             |

| Db         | 144        | CTCAAGGCGGTATACGCTTATCCACAGATACAGGGGATTAACGACGAGAAAGAACATGTG      | 203  |
|------------|------------|---|--|
| QY         | 5277       | agcaaaagccagcaaaaagccsagaaacgltlaaaaagccgcgttgtcgtgctlttcca       | 5336   |
| Db         | 204        | AGCAAAAGCGCCAGCAAAAGGCCAGGAACCGTAAAAAAGCGGCGTTGCGCTTTTTC          | 263  |
| QY         | 5337       | tagatccgcgccttcagagagatcaaaaatcgaagctcaagltcaagatgtgcgaaa         | 5396   |
| Db         | 264        | TAGGCTCCGCCCCCTTACAGCATCATCAAAAATTCACGCTTAAAGTCAAGAGTGGCGAAA      | 323  |
| QY         | 5397       | ccgcacagagactataaagataccagagcglttcccccctggaagctccctcgtgtgcctcc    | 5456   |
| Db         | 324        | CCGACACAGACTATTAAGATATACCAGGGGTTTCCCTCGAGACGTCCTCGTCCGCTTCC       | 383  |
| QY         | 5457       | tgttcgaacctgcgcgtctacgcgatacctgttcgccttcccttccttcgggaagcgtgac     | 5516   |
| Db         | 384        | TGTTTCCACACCTGCGCGCTTACCGGAATACCTGTCGCGCTTTCCTTCGGAAGCGTGGC       | 443  |
| QY         | 5517       | gctttccaatgtccacagctgtatagatcctcaagttcgtgtatagtcgttcgtctccaagct   | 5576   |
| Db         | 444        | GCTTTTCATATAGCTACGCGTGTAGATATCTCACTTCGATGATGATGTCGTTCCGCTCCAACT   | 503  |
| QY         | 5577       | ggcgtcgtgtgaacgaaccccccglttaagccagccgctgtgcgcttalcgcgtlaactacgc   | 5636   |
| Db         | 504        | GGGCTGTGTGACGAGAACCCCCCGTTACGCGCCGACCGCGCTGCGCTTATCCGGTAATATCG    | 563  |
| QY         | 5637       | tcttgatccaacctccggttaagacaagactatctgcacatcggcagaagcgcactgtgtaaacg | 5656   |
| Db         | 564        | TCTTGATGTCACACCCGGGTAGACACGATTTATGCCACATGGGACGACGACACTGGTAACAG    | 623  |
| QY         | 5697       | gattagcagagccgaagatcatgtagcggtgcgtacagagttcttgaagtgtgtgcctaacta   | 5756   |
| Db         | 624        | GATTAGCAGAGCCGAGGATATGTAGGCGGTGTATAGAGGTTCTTGAAAGTGTGTGCCCTAACTA  | 683  |
| QY         | 5757       | cgctcacactagaagacag   | 5776   |
| Db         | 684        | GGGCTACACTAGAGAACAG   | 703  |
| RESULT 11  |            |   |  |
| LOCUS      | A2547271/c |   |  |
| DEFINITION | A2547271   | 887 bp  | DNA  |
| ACCESSION  | ENTMEZ67F  | Entamoeba histolytica   | Sheared DNA  |
| VERSION    | A2547271   | Entamoeba histolytica   | genomic, DNA sequence.   |
| KEYWORDS   | A2547271.1 | GI:11169811   | GSS.   |
| SOURCE     |            |   | Entamoeba histolytica.   |
| ORGANISM   |            |   | Entamoeba histolytica  |
| REFERENCE  |            |   | Eukaryota; Entamoebidae; Entamoeba.                                |
| AUTHORS    |            |   | 1 (bases 1 to 887)   |
| TITLE      |            |   | Loftus, B., Van Aken, S. and Fraser, C.                            |
| JOURNAL    |            |   | Determination of clone end sequences from Entamoeba histolytica    |
| COMMENT    |            |   | HMI:IMSS sheared DNA library                                       |
|            |            |   | Unpublished (2000)   |
|            |            |   | Contact: Brendan J Loftus  |
|            |            |   | Department of Eukaryotic Genomics                                  |
|            |            |   | The Institute for Genomic Research                                 |
|            |            |   | 9712 Medical Center Dr., Rockville, MD 20850, USA                  |
|            |            |   | tel: 301 838 0200  |
|            |            |   | fax: 301 838 0208  |
|            |            |   | Email: b1loftus@tigr.org   |
|            |            |   | Clones are derived from the Entamoeba histolytica HMI:IMSS sheared |
|            |            |   | DNA library  |
|            |            |   | Seq primer: M13-Forward  |
|            |            |   | Class: Shotgun   |
|            |            |   | High quality sequence start: 78                                    |
|            |            |   | High quality sequence stop: 780.                                   |
|            |            |   | Location/Qualifiers  |
|            |            |   | 1..887   |
|            |            |   | /organism="Entamoeba histolytica"                                  |
|            |            |   | /strain="HMI:IMSS"   |



|  |   |  |                      |
|--|---|--|----------------------|
| Db   | 508   | CAAGTCATTCTGAGATATGATGTGGGGGACCGAGTTCTCTTGCCGGCGTCATTAAG         | 567                  |
| Qy   | 6711  | ggataataacgcgcgcacatacagaaacttaaaagtctcatcatcttgtaaaacgtcttc     | 6770                 |
| Db   | 568   | GGATATATACCGGGCCACATAGCAGAACTTTAAAGTCCTCATCATTTGGAAACGTTCTTC     | 627                  |
| Qy   | 6771  | ggggcgaaaacctcgaagatcttcacggtgttgagatcgaatcgaatgtaaccactcg       | 6830                 |
| Db   | 628   | GGGGCGAAACCTCTCAGAGATCTACCGCGCTTGAGATTCGATTCGATGTAAACCACTCG      | 687                  |
| Qy   | 6831  | tgaccaccaactgattcttcagcactc                                      | 6854                 |
| Db   | 688   | TGCACCCCACTGATCTTCAGCATC   | 711                  |
| RESULT 13  |   |  |                      |
| LOCUS  | Bf663156  | 757 bp   | mRNA EST 19-JAN-2001 |
| DEFINITION   | 96304JF03.x1 C. reinhardtii CC-1690, stress condition I, normalized     |  |                      |
| ACCESSION  | Bf663156  |  |                      |
| VERSION  | Bf663156.1  |  |                      |
| KEYWORDS   | EST.  |  |                      |
| SOURCE   | Chlamydomonas reinhardtii.  |  |                      |
| ORGANISM   | Chlamydomonas reinhardtii   |  |                      |
| REFERENCE  | Grossman, A., Davies, J., Federpiel, N., Harris, E., Hauser, C.,        |  |                      |
| AUTHORS  | Lefebvre, P., McDermott, J. P., Shreager, J., Stillow, C. and Stern, D. |  |                      |
| TITLE  | Analyses of the Chlamydomonas reinhardtii genome: A Model,              |  |                      |
| JOURNAL  | Unicellular System for Analyzing Gene Function and Regulation in        |  |                      |
| COMMENT  | Vascular Plants: project phase 3  |  |                      |
|  | Unpublished (2000)  |  |                      |
|  | Contract: Charles Hauser  |  |                      |
|  | DCMB Box 91000  |  |                      |
|  | Duke University   |  |                      |
|  | Durham, NC 27708-1000   |  |                      |
|  | Tel: 919 613 8159   |  |                      |
|  | Fax: 919 613 8177   |  |                      |
|  | Email: chausere@duke.edu.   |  |                      |
| FEATURES   | Location/Qualifiers   |  |                      |
| SOURCE   | 1..757  |  |                      |
|  | /organism="Chlamydomonas reinhardtii"                                   |  |                      |
|  | /strain="CC-1690 wild type mt+ 21gr"                                    |  |                      |
|  | /db_xref="taxon:3055"   |  |                      |
|  | /clone_lib="C. reinhardtii CC-1690, stress condition I,                 |  |                      |
|  | normalized, lambda zap II"  |  |                      |
|  | /note="Vector: pBluescript II SK-; Site.1: EcoRI; Site.2:               |  |                      |
|  | XhoI. This library, constructed by John Davies and Jeffrey              |  |                      |
|  | McDermott, combines cDNAs from CC-1690 cells grown to                   |  |                      |
|  | mid-log phase in TAP-N (30 min, 1hr, 4hr), TAP-S (30 min,               |  |                      |
|  | 1hr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, 1hr              |  |                      |
|  | , 4hr) and NH4 to NO3 (30min, 1hr, 4hr). PolyA mRNA was                 |  |                      |
|  | purified from each sample, pooled and cDNA synthesized.                 |  |                      |
|  | The cDNA was directionally cloned into lambda zap II                    |  |                      |
|  | (Stratagene) in the EcoRI (5') and XhoI (3') sites.                     |  |                      |
|  | pBluescript II SK- plasmids were excised from the lambda                |  |                      |
|  | zap clones by superinfection with Exassist (Stratagene)                 |  |                      |
|  | phage. The library was normalized using method 4 described              |  |                      |
|  | in Bonaldo et al (1996) Genome Research 6: 791-806."                    |  |                      |
| BASE COUNT   | 167 a   | 219 c  | 208 g                |
| ORIGIN   | 161 t   | 2 others   |                      |
| Query Match 8.2%; Score 652.2; DB 170; Length 757;           |   |  |                      |
| Best Local Similarity 98.7%; Pred. No. 8.2e-177;             |   |  |                      |
| Matches 668; Conservative 0; Mismatches 8; Indels 1; Gaps 1; |   |  |                      |
| Qy   | 5091  | ccagctgcattatgaatcgcgcacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc | 5150                 |
| Db   | 75  | CCAGCTGCATTATGAATCGCCCAACGCCCGGGGAGAGCGCGTTGGCTATTGGCGCCTC       | 134                  |

|            |   |  |                         |
|------------|---|--|-------------------------|
| Oy         | 5151  | tttgcttcttcgcttaactgaacagcgtgtgcgcgcgttcgttgtgtgcggagcggtatc           | 5210                    |
| Db         | 135   | TTTCGGTCTTCCTCCACTACATCCTGCGTGCGTGGCTGTGCCTGGCGAGGGATTG                | 194                     |
| Oy         | 5211  | aggccattcaaaagcgjglaalaacgyltaaccacaagaalccaagggataaacgcagaaaaga       | 5270                    |
| Db         | 195   | AGCTCACTCAAAGCGCGGTATACGGTTATCCACAAGATCACGGGGATTAAAGGAAAAA             | 254                     |
| Oy         | 5271  | catgtgagcaaaagcgccagcaaaaagccaggaaccggtataaaagccgcgttgtgcgcgtt         | 5330                    |
| Db         | 255   | CATGTGAGCAAAAAGGCCAGCAAAAAGGCCAGGAACCGTAAAAAGGCCCGCTTGCCTGGCGTT        | 314                     |
| Oy         | 5331  | tttccataagctcgcgcccccctgaagagatatcaaaaaaatltgaagctcaagttaagaggtc       | 5390                    |
| Db         | 315   | TTTTCATATAGCTCGCCCCCTTAGCAGCATACAAAATATGACGCTCAAGTCAGAGGTG             | 374                     |
| Oy         | 5391  | gcgaaacccgcacagagactaaagaatacaccaagcgcttlccccctlbgaaagctccctcgtgcg     | 5450                    |
| Db         | 375   | GCGAACC CGACA GGA ACTATAA GAATAC CAGGCG GTTCC CCC CTG AAG CTC CC CGCGG | 434                     |
| Oy         | 5451  | cctccctgtgtccgaacctgcgcgttacacgataacctgtgcgccttctcccttcgggaag          | 5510                    |
| Db         | 435   | CTCTCCCTGTTTCCAGACCCCTGCGCGTTACCGGAACTCTGCGCCCTTTCCCTTCGGGAAAG         | 494                     |
| Oy         | 5511  | cgtgcgagctttcccaatgctcaacgcgtgaagatctcaagttcgcgttgtgaagttgtgcgcc       | 5570                    |
| Db         | 495   | CGTGGCCCTTCTCATATAGCTCAAGCTGTAGATCTCATGTTCCGGTGTAGGTGCTTGCGCTC         | 554                     |
| Oy         | 5571  | caagcttgcgctgtgtgcacgaaccccccgctttaagccagcaaccgcgtgcgccttatccgylaa     | 5630                    |
| Db         | 555   | CAAGCTGGGCTGTGTGCACGAACCCCCGTTGACGCCGACCGCTTATTCGGTA                   | 614                     |
| Oy         | 5631  | ctatcgtccttgaatccaaccgcgttaagaacagactatcgcgcactatgcaacagccactcg        | 5690                    |
| Db         | 615   | CTATCGCTTTGATTCACACCCCGTAAAGACACACTTATCGCACTGGCACACCACTGG              | 674                     |
| Oy         | 5691  | taacagatttagcaagcgaggtatgatgtgaagtgctcacaagttcttgaagtgtagcc            | 5750                    |
| Db         | 675   | TAAACAGATTATACGACAGACCGAGATGTATGACCGCGTCTACAGAATCTTTGGAGTAGATAGC       | 734                     |
| Oy         | 5751  | taactacgctactacta  | 5767                    |
| Db         | 735   | TAAAC-ACGGCTACACTA   | 750                     |
| RESULT_14  | AL040542/c  |  |                         |
| LOCUS      | AL040542  | 767 bp   | mRNA                    |
| DEFINITION | DKFZP434I1114.J1  | 434 (synonym: htes3)   | Homo sapiens cDNA clone |
| VERSION    | AL040542  |  |                         |
| KEYWORDS   | AL040542.1  | GI:5409490   |                         |
| SOURCE     | EST.  |  |                         |
| ORGANISM   | Homo sapiens  |  |                         |
| REFERENCE  | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;   |  |                         |
| AUTHORS    | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.           |  |                         |
| TITLE      | 1 (bases 1 to 767)  |  |                         |
| JOURNAL    | Koehler,K., Beyer,A., Mewes,H.W., Gassenhuber,J. and Wiemann,S.     |  |                         |
| COMMENT    | EST (Koehler, et al.)   |  |                         |
|            | Unpublished (1999)  |  |                         |
|            | Contact: Koehler K  |  |                         |
|            | MTS   |  |                         |
|            | Am Kioferspitze 18a D-82152 Martinsried, Germany                    |  |                         |
|            | This is the 5' sequence of the clone Insert                         |  |                         |
|            | Clone from S. Wiemann, Molecular Genome Analysis, German Cancer     |  |                         |
|            | Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;         |  |                         |
|            | sequenced by BMFZ (Biomedical Research Center at the Charite,       |  |                         |
|            | Berlin/Germany) within the cDNA sequencing consortium of the German |  |                         |
|            | Genome Project.   |  |                         |
|            | No sl sequence available.   |  |                         |
|            | This clone (DKFZP434I114) is available at the R2PD in Berlin.       |  |                         |





Db 198 TTAAGAAGTGCTCTTCATTTGGAAANNCTGTTCTTTCGGCCGGAANCTNCNMAAGATCTTAC 139

Qy 6797 GCGTGTGAGATCCAACTGATGATGAACCACTGTCGACCAACTGATCTTCACATCTT 6850

Db 138 CCGTGATGAGATCCAGTTCATGTAAACCCACTGTCGACCAACTGATCTTCACATCTT 79

Qy 6857 TCACTTCAACAGCGTTCCTGATGAGCAAAACAAGGCAAAATGCGCG 6908

Db 78 TTACTTTCACACGCGTTTCNGGGGNCNCAAAACAGGAGGCAAAATGCCCTC 27

Search completed: November 18, 2001, 04:30:13  
Job time: 39895 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 17, 2001, 19:06:38 ; Search time 389.82 Seconds  
(without alignments)  
4615.313 Million cell updates/sec

Title: US-09-689-430-1

Perfect score: 7944  
1 tggccactctctctctgcgc.....aggggggggggggggggggt 7944

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/prodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/PCTUS.COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/backfile1.seq:\*

Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | DB | ID                | Description       |
|------------|--------|-------------|--------|----|-------------------|-------------------|
| 1          | 6608.4 | 83.2        | 11933  | 4  | US-09-470-618-13  | Sequence 13, Appl |
| 2          | 6608.4 | 83.2        | 11933  | 4  | US-09-364-862-13  | Sequence 13, Appl |
| 3          | 4456.2 | 56.1        | 4999   | 4  | US-09-470-618-14  | Sequence 14, Appl |
| 4          | 4456.2 | 56.1        | 4999   | 4  | US-09-364-862-14  | Sequence 14, Appl |
| 5          | 4413   | 55.6        | 4629   | 2  | US-08-484-891-7   | Sequence 7, Appl  |
| 6          | 4411.6 | 54.9        | 9354   | 1  | US-08-683-839B-2  | Sequence 2, Appl  |
| 7          | 4357.6 | 47.2        | 4670   | 3  | US-08-717-294-41  | Sequence 41, Appl |
| 8          | 3748   | 47.2        | 5035   | 2  | US-08-882-083-1   | Sequence 1, Appl  |
| 9          | 3748   | 47.2        | 5035   | 2  | US-08-558-107-1   | Sequence 1, Appl  |
| 10         | 3748   | 47.2        | 5035   | 3  | US-09-243-539-1   | Sequence 1, Appl  |
| 11         | 3189.8 | 40.2        | 4334   | 2  | US-08-670-707A-38 | Sequence 38, Appl |
| 12         | 3189.8 | 40.2        | 4334   | 4  | US-09-037-601-38  | Sequence 38, Appl |
| 13         | 2985.2 | 37.6        | 6253   | 4  | US-08-893-327-15  | Sequence 15, Appl |
| 14         | 2985.2 | 37.6        | 6280   | 3  | US-08-893-327-17  | Sequence 17, Appl |
| 15         | 2985.2 | 37.6        | 6280   | 3  | US-08-893-327-19  | Sequence 19, Appl |
| 16         | 2758.8 | 34.7        | 3984   | 1  | US-08-040-753-1   | Sequence 1, Appl  |
| 17         | 2656.2 | 33.2        | 4696   | 2  | US-08-929-967-15  | Sequence 15, Appl |
| 18         | 2656.2 | 33.2        | 5158   | 2  | US-08-929-967-16  | Sequence 16, Appl |
| 19         | 2656.2 | 33.2        | 7607   | 1  | US-08-222-616-19  | Sequence 19, Appl |
| 20         | 2656.2 | 33.2        | 7607   | 5  | PCT-US95-04228-19 | Sequence 19, Appl |
| 21         | 2636.2 | 33.2        | 9108   | 5  | PCT-US95-04228-45 | Sequence 45, Appl |
| 22         | 2636.2 | 33.2        | 19307  | 5  | US-08-836-022A-10 | Sequence 10, Appl |
| 23         | 2636.2 | 33.2        | 19307  | 4  | US-09-427-048A-10 | Sequence 10, Appl |
| 24         | 2625.8 | 33.1        | 4054   | 3  | US-09-098-287A-9  | Sequence 9, Appl  |
| 25         | 2605.4 | 32.8        | 4451   | 3  | US-08-717-294-42  | Sequence 42, Appl |
| 26         | 2593.2 | 32.6        | 4454   | 1  | US-07-712-284-1   | Sequence 1, Appl  |
| 27         | 2593.2 | 32.6        | 4454   | 5  | PCT-US92-04227-1  | Sequence 1, Appl  |

|      |        |      |      |   |                  |                  |
|------|--------|------|------|---|------------------|------------------|
| C 28 | 2593.2 | 32.6 | 4732 | 1 | US-07-884-811-1  | Sequence 1, Appl |
| C 29 | 2593.2 | 32.6 | 4732 | 1 | US-07-885-971-1  | Sequence 1, Appl |
| C 30 | 2593.2 | 32.6 | 4732 | 1 | US-08-087-783A-1 | Sequence 1, Appl |
| C 31 | 2593.2 | 32.6 | 4732 | 1 | US-08-194-088B-1 | Sequence 1, Appl |
| C 32 | 2593.2 | 32.6 | 4732 | 2 | US-08-194-087-1  | Sequence 1, Appl |
| C 33 | 2593.2 | 32.6 | 4732 | 5 | PCT-US93-04648-1 | Sequence 1, Appl |
| C 34 | 2593.2 | 32.6 | 5141 | 1 | US-08-286-305A-9 | Sequence 9, Appl |
| C 35 | 2593.2 | 32.6 | 5141 | 2 | US-08-441-104A-9 | Sequence 9, Appl |
| C 36 | 2593.2 | 32.6 | 5141 | 4 | US-08-440-816A-9 | Sequence 9, Appl |
| C 37 | 2593.2 | 32.6 | 5141 | 4 | US-09-417-381A-9 | Sequence 9, Appl |
| C 38 | 2593.2 | 32.6 | 6557 | 1 | US-08-286-740-3  | Sequence 3, Appl |
| C 39 | 2593.2 | 32.6 | 6557 | 5 | PCT-US95-09576-3 | Sequence 3, Appl |
| C 40 | 2593.2 | 32.6 | 6889 | 1 | US-08-286-740-2  | Sequence 2, Appl |
| C 41 | 2593.2 | 32.6 | 6889 | 5 | PCT-US95-09576-2 | Sequence 2, Appl |
| C 42 | 2593.2 | 32.6 | 7305 | 5 | US-08-286-740-4  | Sequence 4, Appl |
| C 43 | 2593.2 | 32.6 | 7305 | 5 | PCT-US95-09576-4 | Sequence 4, Appl |
| C 44 | 2593.2 | 32.6 | 7360 | 1 | US-08-286-740-1  | Sequence 1, Appl |
| C 45 | 2593.2 | 32.6 | 7360 | 5 | PCT-US95-09576-1 | Sequence 1, Appl |

## ALIGNMENTS

```
RESULT 1
US-09-470-618-13
; Sequence 13, Application US/09470618
; Patent No. 6200560
; GENERAL INFORMATION:
; APPLICANT: Couto, Linda B.
; APPLICANT: Colosl, Peter C.
; TITLE OF INVENTION: Adeno-Associated Vectors for Expression of Factor VIII
; TITLE OF INVENTION: by Target Cells
; FILE REFERENCE: Avigen-04082
; CURRENT APPLICATION NUMBER: US/09/470,618
; CURRENT FILING DATE: 1999-12-22
; EARLIER APPLICATION NUMBER: 09/364,862
; EARLIER FILING DATE: 1999-07-30
; EARLIER APPLICATION NUMBER: 60/125,974
; EARLIER FILING DATE: 1999-03-24
; EARLIER APPLICATION NUMBER: 60/104,994
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentlin Ver. 2.0
; SEQ ID NO 13
; LENGTH: 11933
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-470-618-13

Query Match      83.2%; Score 6608.4; DB 4; Length 11933;
Best Local Similarity 91.9%; Pred. No. 0;
Matches 7211; Conservative 0; Mismatches 256; Indels 379; Gaps 7;

QY 15 ctggcgctgcgtgcctcactgaagccggcgagcaagaagtcgccgagccgggctt 74
      |||||||
DB 4 ctggcgctgcgtgcctcactgaagccggcgagcaagaagtcgccgagccgggctt 63
      |||||||
QY 75 gccggcgagccctcagtgagcgagcgagcgagcgaggaagtggtgccaactcact 134
      |||||||
DB 64 gttccgcgcgcctcagtgagcgagcgagcgagcgaggaagtggtgccaactcact 123
      |||||||
QY 135 aggggttctt-----cagatctcttctaagtaaacagtaacatgaacctt 180
      |||||||
DB 124 aggggttctcggcgagcgagcgaggaagtggttcttctaagtaaacagtaacatgaacctt 183
      |||||||
QY 181 accccgtgtcgtggcaagcgccgtgtgtgccaagtggttgcgtgaagcaacccacctg 240
      |||||||
DB 184 ttcttaataaccatccaggaatgttgttcttctaataaccatccagtaagtattgttaaa 243
      |||||||
QY 241 gctgggcttcggcatagcgcatcagcgca-tgcgatactcagtggttgcgaagag 299
```

|    |      |   |      |
|----|------|---|------|
| Db | 244  | gaagtatattaaagcgaagctcttctctgcacacagatacacttccgggtgcccgcctctgg          | 303  |
| Qy | 300  | aagcaaaaagacctctccaccccaagccttggaatgtttccaccccaatgttcgaagatgttgct       | 359  |
| Db | 304  | caagtaagtccgtgtgtgttccgcgggcctgttaccggtgtatagcccttgc                    | 363  |
| Qy | 360  | tttgcaagagggaagcaaaaagcctctccaccccaagccctcgagagct--tcgaac               | 416  |
| Db | 364  | gtgccttgaaattccgacatgcataccacttttcttctccacaggtatcgatcc                  | 423  |
| Qy | 417  | accattgcaaatagaagctctccacactgtctcttctgtgccttttgcaatttcgtcttagt          | 476  |
| Db | 424  | accattgcaaatagaagctctccacactgtctcttctgtgccttttgcaatttcgtcttagt          | 483  |
| Qy | 477  | gccaccaagaatactactactgtggtgtgcagtggaactgttcatatggagctatattgcaaatgtat    | 536  |
| Db | 484  | gccaccaagaagaatactactactgtggtgtgcagtggaactgttcatatggagctatattgcaaatgtat | 543  |
| Qy | 537  | ctcgggtgaagctgtcgtgtggaagcgaagaattccctccctagaagtgccaaaattcttccatcc      | 596  |
| Db | 544  | ctcgggtgaagctgtcgtgtggaagcgaagaattccctccctagaagtgccaaaattcttccatcc      | 603  |
| Qy | 597  | aacacctcagtcgtgtgtaacaaaagactcgttltgttagaattcacggttccacttttccac         | 656  |
| Db | 604  | aacacctcagtcgtgtgtaacaaaagactcgttltgttagaattcacggttccacttttccac         | 663  |
| Qy | 657  | atcgcttaagccaagggccacccctggaatggtctgtctagaattccataccatccagcctggaagt     | 716  |
| Db | 664  | atcgcttaagccaagggccacccctggaatggtctgtctagaattccataccatccagcctggaagt     | 723  |
| Qy | 717  | tatgatacagtggttcatattacacttaagaacaatggcttcccatccctgtacgttcatatgct       | 776  |
| Db | 724  | tatgatacagtggttcatattacacttaagaacaatggcttcccatccctgtacgttcatatgct       | 783  |
| Qy | 777  | gttgggtgatccctcctacatggaagctctctcgagggaagctgaatatgtatgtacacacagttcaa    | 836  |
| Db | 784  | gttgggtgatccctcctacatggaagctctctcgagggaagctgaatatgtatgtacacacagttcaa    | 843  |
| Qy | 837  | agggaagaaaagaaatgaataagcttccctgtgtgaaagccaataatgtcttgcagagtc            | 896  |
| Db | 844  | agggaagaaaagaaatgaataagcttccctgtgtgaaagccaataatgtcttgcagagtc            | 903  |
| Qy | 897  | ctgaaagaaagaaatgttccaatgtgctctgcacccactgtgccttaacctactaatactttct        | 956  |
| Db | 904  | ctgaaagaaagaaatgttccaatgtgctctgcacccactgtgccttaacctactaatactttct        | 963  |
| Qy | 957  | catgttggaagcttgtaaaagacttgaattcgaagcctcatatggagccctactatgatatgata       | 1016 |
| Db | 964  | catgttggaagcttgtaaaagacttgaattcgaagcctcatatggagccctactatgatatgata       | 1023 |
| Qy | 1017 | gaagggagcttcggccaaagaaaagacacagaccttgcacaaattatatactatttttgt            | 1076 |
| Db | 1024 | gaagggagcttcggccaaagaaaagacacagaccttgcacaaattatatactatttttgt            | 1083 |
| Qy | 1077 | gtaatttgatgaagggaagaatttgcacacttcgnaaaacaaagaactcctctgtatgcagatagg      | 1136 |
| Db | 1084 | gtaatttgatgaagggaagaatttgcacacttcgnaaaacaaagaactcctctgtatgcagatagg      | 1143 |
| Qy | 1137 | gatgcgcacatctcgtccgggcttgcccaaaatgcacacagccaatgtgtatglaaaacgg           | 1196 |
| Db | 1144 | gatgcgcacatctcgtccgggcttgcccaaaatgcacacagccaatgtgtatglaaaacgg           | 1203 |
| Qy | 1197 | tctctgcacaggtctgatgttgatgcacacaggaataatcagatctatattggcatgtgatgtgaaatg   | 1256 |
| Db | 1204 | tctctgcacaggtctgatgttgatgcacacaggaataatcagatctatattggcatgtgatgtgaaatg   | 1263 |
| Qy | 1257 | ggacacacttcgaagatgcacactaaataattctctgaaaggtcacacattctctgttgaagac        | 1316 |
| Db | 1264 | ggacacacttcgaagatgcacactaaataattctctgaaaggtcacacattctctgttgaagac        | 1323 |
| Qy | 1317 | catcgcagagctgcctcttggaatatctgcgcaataacttctcttactgtccaaacactctgt         | 1376 |
| Db | 1324 | catcgcagagctgcctcttggaatatctgcgcaataacttctcttactgtccaaacactctgt         | 1383 |
| Qy | 1377 | atggaacttggacagttcttctactgttctgtgtaalcttccccaacacatgatgtgcatg           | 1436 |
| Db | 1384 | atggaacttggacagttcttctactgttctgtgtaalcttccccaacacatgatgtgcatg           | 1443 |
| Qy | 1437 | gaagcttatgtcaaaagtgaagagctgtctcagaagggaaaccccaactacagatagaataaat        | 1496 |
| Db | 1444 | gaagcttatgtcaaaagtgaagagctgtctcagaagggaaaccccaactacagatagaataaat        | 1503 |
| Qy | 1497 | gaagaagcgggaagactatgatgatgatcttactgtattctgtgaatgtatgtgtcaagttt          | 1556 |
| Db | 1504 | gaagaagcgggaagactatgatgatgatcttactgtattctgtgaatgtatgtgtcaagttt          | 1563 |
| Qy | 1557 | gatgtatgaacatctcttcccttcttccaaatctgctaagtttccagaagaatccctcaa            | 1616 |
| Db | 1564 | gatgtatgaacatctcttcccttcttccaaatctgctaagtttccagaagaatccctcaa            | 1623 |
| Qy | 1617 | acttggtatcatattacatgtctgtctaagaagggagactggagactatgctcccttactgctc        | 1676 |
| Db | 1624 | acttggtatcatattacatgtctgtctaagaagggagactggagactatgctcccttactgctc        | 1683 |
| Qy | 1677 | ggccccgatgacagaagaattataaaagtcaataatttgaacaaatggtccctccagcgatgtgt       | 1736 |
| Db | 1684 | ggccccgatgacagaagaattataaaagtcaataatttgaacaaatggtccctccagcgatgtgt       | 1743 |
| Qy | 1737 | aggaagtaacaaaaggtccgatattatgtcatacaagaatgtgaaccttaagaatccgttaa          | 1796 |
| Db | 1744 | aggaagtaacaaaaggtccgatattatgtcatacaagaatgtgaaccttaagaatccgttaa          | 1803 |
| Qy | 1797 | gctattcagcatgaaatcgaagaacttggagccttacttacttattatgggaaagtttgagacaca      | 1856 |
| Db | 1804 | gctattcagcatgaaatcgaagaacttggagccttacttacttattatgggaaagtttgagacaca      | 1863 |
| Qy | 1857 | ctgttgattatataatttaagaataacagaagcagacataatacatctaccctccaggaatc          | 1916 |
| Db | 1864 | ctgttgattatataatttaagaataacagaagcagacataatacatctaccctccaggaatc          | 1923 |
| Qy | 1917 | actgcatctgcgtccttcttattcgaagggatattcccaaaaaggttgttaaaacatttgaagat       | 1976 |
| Db | 1924 | actgcatctgcgtccttcttattcgaagggatattcccaaaaaggttgttaaaacatttgaagat       | 1983 |
| Qy | 1977 | tttccaattctgcgcgggagaatatattcaataataatgtgaacagtgactgttgaagatggg         | 2036 |
| Db | 1984 | tttccaattctgcgcgggagaatatattcaataataatgtgaacagtgactgttgaagatggg         | 2043 |
| Qy | 2037 | ccaactaatcagatccctcgtgtgccttggaacccgtatactactagtttgcgttaataatgag        | 2096 |
| Db | 2044 | ccaactaatcagatccctcgtgtgccttggaacccgtatactactagtttgcgttaataatgag        | 2103 |
| Qy | 2097 | agagattctagcttcagagactatggccctctctctcatctcgtctatacagaagaatctgttagat     | 2156 |
| Db | 2104 | agagattctagcttcagagactatggccctctctctcatctcgtctatacagaagaatctgttagat     | 2163 |
| Qy | 2157 | caaagaggaaacagataatgtgcagacaagaaggaaatgtcatccgtgttcttcgtatattgat        | 2216 |
| Db | 2164 | caaagaggaaacagataatgtgcagacaagaaggaaatgtcatccgtgttcttcgtatattgat        | 2223 |
| Qy | 2217 | gagaacccaagactgtgtactccaagaagataatacagcttcttcccaatccagcttga             | 2276 |
| Db | 2224 | gagaacccaagactgtgtactccaagaagataatacagcttcttcccaatccagcttga             | 2283 |
| Qy | 2277 | gtgcagcttggaagatccaagattccaagcctccaacatcatatgcacagcataatgtgctat         | 2336 |
| Db | 2284 | gtgcagcttggaagatccaagattccaagcctccaacatcatatgcacagcataatgtgctat         | 2343 |
| Qy | 2337 | gtttttgtagtttgcaattgttgcagttgttttgcaatggaggtggcaatgttataattcta          | 2396 |
| Db | 2344 | gtttttgtagtttgcaattgttgcagttgttttgcaatggaggtggcaatgttataattcta          | 2403 |
| Qy | 2397 | agcatttgagacaagaactgcacttccctctcgtctctctctctgtgatalaacttccaaac          | 2456 |
| Db | 2404 | agcatttgagacaagaactgcacttccctctcgtctctctctctgtgatalaacttccaaac          | 2463 |

|    |      |   |      |
|----|------|---|------|
| OY | 2457 | aaatggtctataagaacacataccatactcattccattccattccaggaagaactgcttcatg       | 2516 |
| Db | 2464 | aaaaatgcttatataagaacacataccataccattccattccattccaggaagaactgcttcatg     | 2523 |
| OY | 2517 | tcgaatggaataaaccaagctctatagatctctcgtgggtgcacaaactcagaacttcggaaacaga   | 2576 |
| Db | 2524 | tcgattggaataaaccaagctctatagatctctcgtgggtgcacaaactcagaacttcggaaacaga   | 2583 |
| OY | 2577 | ggcatgacccgctctactgtaagaaggtttcttagttgtgacaaagaacactgctgattatcagag    | 2636 |
| Db | 2584 | ggcatgacccgctctactgtaagaaggtttcttagttgtgacaaagaacactgctgattatcagag    | 2643 |
| OY | 2637 | gaacggttatataagaatatttccagatactcttcgtcgatgaataaacaatgcatattgaaaccaaga | 2696 |
| Db | 2644 | gaacggttatataagaatatttccagatactcttcgtcgatgaataaacaatgcatattgaaaccaaga | 2703 |
| OY | 2697 | agctctctcccaagatccaagaaccccttagcactagcgaagcgaagcaaatlaagccaacca       | 2756 |
| Db | 2704 | agctctctcccaagatccaagaaccccttagcactagcgaagcgaagcaaatlaagccaacca       | 2709 |
| OY | 2757 | ccagatcttgaaacgcatacaacggaataactcgttactactcttcagtcagatcaagaag         | 2816 |
| Db | 2710 | -----gaaataactcgttactactcttcagtcagatcaagaag                           | 2745 |
| OY | 2817 | gaaatctactatgatataccatactcagttgaatgaagaagaagaagatcttgacatttat         | 2876 |
| Db | 2746 | gaaatctactatgatataccatactcagttgaatgaagaagaagaagatcttgacatttat         | 2805 |
| OY | 2877 | gattgagatgaaaaatcaagaccccccgaagcttccaaaagaacacgaacattatttat           | 2936 |
| Db | 2806 | gattgagatgaaaaatcaagaccccccgaagcttccaaaagaacacgaacattatttat           | 2865 |
| OY | 2937 | gctgcagatggagaaggtctcggggaattatggagatgaatctcccaacatggttctaagaac       | 2996 |
| Db | 2866 | gctgcagatggagaaggtctcggggaattatggagatgaatctcccaacatggttctaagaac       | 2925 |
| OY | 2997 | aggggtccaagatggacagttccctccagttccaagaagaagttgtttccagaagaattacagat     | 3056 |
| Db | 2926 | aggggtccaagatggacagttccctccagttccaagaagaagttgtttccagaagaattacagat     | 2985 |
| OY | 3057 | ggctcccttctactcagcccttataccgttggagaacccaatgaacatttggagactccgtggg      | 3116 |
| Db | 2986 | ggctcccttctactcagcccttataccgttggagaacccaatgaacatttggagactccgtggg      | 3045 |
| OY | 3117 | ccatataaagagcagaagtttgaagaataatcatcagtgtaacttccaagaatacagccctc        | 3176 |
| Db | 3046 | ccatataaagagcagaagtttgaagaataatcatcagtgtaacttccaagaatacagccctc        | 3105 |
| OY | 3177 | cgctccattactccttactatctagccttattcttctatgsgaagatcagagggcaaggagca       | 3236 |
| Db | 3106 | cgctccattactccttactatctagccttattcttctatgsgaagatcagagggcaaggagca       | 3165 |
| OY | 3237 | gaaactggaanaaaactttgtcaagcctaaatgaacaaaacttcttttgaagaagtgcga          | 3296 |
| Db | 3166 | gaaactggaanaaaactttgtcaagcctaaatgaacaaaacttcttttgaagaagtgcga          | 3225 |
| OY | 3297 | catcatatggcacccctcctaanaatgatgttgcagctgcgaagccttgaggtctatttcccgat     | 3356 |
| Db | 3226 | catcatatggcacccctcctaanaatgatgttgcagctgcgaagccttgaggtctatttcccgat     | 3285 |
| OY | 3357 | gttggaactggaaaaaagaatgtgcaactcgaagcttgatattggaacccctctggtctgcacact    | 3416 |
| Db | 3286 | gttggaactggaaaaaagaatgtgcaactcgaagcttgatattggaacccctctggtctgcacact    | 3345 |
| OY | 3417 | aaacacatggaacccctctcagttgggagacaagatggaacagatggagattggccgttttttc      | 3476 |
| Db | 3346 | aaacacatggaacccctctcagttgggagacaagatggaacagatggagattggccgttttttc      | 3405 |
| OY | 3477 | acaactcttgaagagacaacaaagctgtgtaacttcaactggaataatgatggaagaacatgcag     | 3536 |
| Db | 3406 | acaactcttgaagagacaacaaagctgtgtaacttcaactggaataatgatggaagaacatgcag     | 3465 |

|    |      |   |      |
|----|------|---|------|
| OY | 3537 | gcccctgcgaatcatcagatcgtgagaagaccacatttaaagaaattatcgtctccatgca           | 3536 |
| Db | 3466 | gtcccttcgcaatcatccagatcgtgagaagatccacatttaaagaaattatcgtctccatca         | 3525 |
| OY | 3597 | atcaatgcctcatcaatcgaatgggtatcacactccgtgttgatgaatggctccagatccaagagt      | 3656 |
| Db | 3526 | atcaatgctctcatcaatcgaatgggtatcacactccgtgttgatgaatggctccagatccaagagt     | 3555 |
| OY | 3657 | cgatcgtatcctcgtccagatcgtggcagcaaatgaaaaatccatctcatcttcaagtga            | 3716 |
| Db | 3586 | cgatcgtatcctcgtccagatcgtggcagcaaatgaaaaatccatctcatcttcaagtga            | 3645 |
| OY | 3717 | catgtgttcaactgtcagcaaaaaaaggagggtataaaatggccgttataatctctatcca           | 3776 |
| Db | 3646 | catgtgttcaactgtcagcaaaaaaaggagggtataaaatggccgttataatctctatcca           | 3705 |
| OY | 3777 | gggtgtttttgagaagatcgtgaaatcgtttaccatcccaaaagctcgtgaatttggcgttggaatgc    | 3836 |
| Db | 3706 | gggtgtttttgagaagatcgtgaaatcgtttaccatcccaaaagctcgtgaatttggcgttggaatgc    | 3765 |
| OY | 3837 | cttatcttgcgagcatctcaactcgtctgggaatggagacacatttctcgtgttatacgaatgaag      | 3886 |
| Db | 3766 | cttatcttgcgagcatctcaactcgtctgggaatggagacacatttctcgtgttatacgaatgaag      | 3825 |
| OY | 3897 | tgtaagactcccccgtgggaatcgtctcttggaacaaattagaagatttcagattacaagcttca       | 3956 |
| Db | 3826 | tgtaagactcccccgtgggaatcgtctcttggaacaaattagaagatttcagattacaagcttca       | 3885 |
| OY | 3957 | ggacaataatggaacagtgtggccccaagctcgtgccagaatcattatctccgataccaat           | 4016 |
| Db | 3886 | ggacaataatggaacagtgtggccccaagctcgtgccagaatcattatctccgataccaat           | 3945 |
| OY | 4017 | ggccgtggagacccaaggagggcccttctcttggaatcaagatggatctctttggccaagatgt        | 4076 |
| Db | 3946 | ggccgtggagacccaaggagggcccttctcttggaatcaagatggatctctttggccaagatgt        | 4005 |
| OY | 4077 | attcaacgcatcacaagaccccagggtggcccgatcagaagattctccagagcttacaatctccag      | 4136 |
| Db | 4006 | attcaacgcatcacaagaccccagggtggcccgatcagaagattctccagagcttacaatctccag      | 4065 |
| OY | 4137 | ttttatcatctgtatagttctgtatgtggagaagatgtgcagacttaccgagggaattccaact        | 4196 |
| Db | 4066 | ttttatcatctgtatagttctgtatgtggagaagatgtgcagacttaccgagggaattccaact        | 4125 |
| OY | 4197 | ggaacccatlaatggtctctctcttcttggaatcgtggaatcatcctcgtgataaaacaacaattttc    | 4256 |
| Db | 4126 | ggaacccatlaatggtctctctcttcttggaatcgtggaatcatcctcgtgataaaacaacaattttc    | 4185 |
| OY | 4257 | aaacctccaattattgtctcgatcatcatccgatttggacaaccaatcatatataagatctcgacg      | 4316 |
| Db | 4186 | aaacctccaattattgtctcgatcatcatccgatttggacaaccaatcatatataagatctcgacg      | 4245 |
| OY | 4317 | acctcttcgcatctgagttgtagtgggctgttgatttaaatagtttgagaatgcgaatttggagatg     | 4376 |
| Db | 4246 | acctcttcgcatctgagttgtagtgggctgttgatttaaatagtttgagaatgcgaatttggagatg     | 4305 |
| OY | 4377 | gagagtaaaagcaatatacagaatgcagaaattactgtctcatcacttaattacaataatgtt         | 4436 |
| Db | 4306 | gagagtaaaagcaatatacagaatgcagaaattactgtctcatcacttaattacaataatgtt         | 4365 |
| OY | 4437 | ggcacccctggtctctctccaagaagcttcgaattcaacccccaaggaggaatgaatgctctggaga     | 4496 |
| Db | 4366 | ggcacccctggtctctctccaagaagcttcgaattcaacccccaaggaggaatgaatgctctggaga     | 4425 |
| OY | 4497 | ccctcaagggtgaataatccaagaaggttggtcgtcgaagatggagactctccagaagacaatgaagaatc | 4556 |
| Db | 4426 | ccctcaagggtgaataatccaagaaggttggtcgtcgaagatggagactctccagaagacaatgaagaatc | 4485 |
| OY | 4557 | acagagatcaactactcaagggagttaaaattctcgtcttaccagcaatgtaatgtgaagagttc       | 4616 |
| Db | 4486 | acagagatcaactactcaagggagttaaaattctcgtcttaccagcaatgtaatgtgaagagttc       | 4545 |
| OY | 4617 | ctcaatctccagcaagatcaagaatgtggccatctgaatgtgaactctctttttccaatcgtgcaagatga | 4676 |

[illegible]





|    |      |  |      |
|----|------|--|------|
| QY | 657  | atcgcctaagcccaagcgccacccttgatggtctgtctaagtctctaacalccagcgctgaagtt      | 716  |
| Db | 664  | atcgcttaagcccaagcgccacccttgatggtggtctgtctaagtctctaacalccagcgctgaagtt   | 723  |
| QY | 717  | tatgatacagttgtctaatctaaccttaagaacatggtctcccaactctgtccagcttccagtt       | 776  |
| Db | 724  | tatgatacagttgtctaatctaaccttaagaacatggtctcccaactctgtccagcttccagtt       | 783  |
| QY | 777  | gttctgtctatccctactctggaagaagctctctgagagagcttgataatgatatgataccagttca    | 836  |
| Db | 784  | gttctgtctatccctactctggaagaagctctctgagagagcttgataatgatatgataccagttca    | 843  |
| QY | 837  | agggagaagaagaatgataaagttctccctcggttgagaagccaataatgctctgtaggtc          | 896  |
| Db | 844  | agggagaagaagaatgataaagttctccctcggttgagaagccaataatgctctgtaggtc          | 903  |
| QY | 897  | ctgaaagaagaatgctccaatggtcctctgaccccaattggtcccttaactactcaactctctc       | 956  |
| Db | 904  | ctgaaagaagaatgctccaatggtcctctgaccccaattggtcccttaactactcaactctctc       | 963  |
| QY | 957  | catctgtagaccctgtaaaagaacttgaaatctcaggtccatacttgagagccctactagtatgtaga   | 1016 |
| Db | 964  | catctgtagaccctgtaaaagaacttgaaatctcaggtccatacttgagagccctactagtatgtaga   | 1023 |
| QY | 1017 | gaaggaagttctggtccaaagaaagacaagaaaccttgcaacaattatactacttttgcct          | 1076 |
| Db | 1024 | gaaggaagttctggtccaaagaaagacaagaaaccttgcaacaattatactacttttgcct          | 1083 |
| QY | 1077 | gatatgtgataagaaggaanaatttgacatctcagaaacaaagaacctcttgatgtagagatag       | 1136 |
| Db | 1084 | gatatgtgataagaaggaanaatttgacatctcagaaacaaagaacctcttgatgtagagatag       | 1143 |
| QY | 1137 | gatctgcgcatctctcgtcgcccttgccctaaanaatgcacacagttcaatggttlatgtlaaacaag   | 1196 |
| Db | 1144 | gatctgcgcatctctcgtcgcccttgccctaaanaatgcacacagttcaatggttlatgtlaaacaag   | 1203 |
| QY | 1197 | tctctgtccaggtcttgatgtggaatgcccacagaagaatcagttactatgtagatgtagattggaatg  | 1256 |
| Db | 1204 | tctctgtccaggtcttgatgtggaatgcccacagaagaatcagttactatgtagatgtagattggaatg  | 1263 |
| QY | 1257 | ggccaccacctctgaagtgacactaaatctctctgaaggtcaacaaattctgtgtagaagac         | 1316 |
| Db | 1264 | ggccaccacctctgaagtgacactaaatctctctgaaggtcaacaaattctgtgtagaagac         | 1323 |
| QY | 1317 | catctgcagagcgctcctctgtaaatctctgcacaaataacttctcttactgtctcaacaacactctg   | 1376 |
| Db | 1324 | catctgcagagcgctcctctgtaaatctctgcacaaataacttctcttactgtctcaacaacactctg   | 1383 |
| QY | 1377 | atggacccttgagacagttcttactgttttgttcatactctctcccaacaabgatgtagcatg        | 1436 |
| Db | 1384 | atggacccttgagacagttcttactgttttgttcatactctctcccaacaabgatgtagcatg        | 1443 |
| QY | 1437 | gaagcttattgtaaaagtatgacagctctgcccagaggaaccccacactagaatgaaanaataat      | 1496 |
| Db | 1444 | gaagcttattgtaaaagtatgacagctctgcccagaggaaccccacactagaatgaaanaataat      | 1503 |
| QY | 1497 | gaagaagcggaagaactatgatagtatcttacttaacgtatctgaatggaatggtgtcaagttt       | 1556 |
| Db | 1504 | gaagaagcggaagaactatgatagtatcttacttaacgtatctgaatggaatggtgtcaagttt       | 1563 |
| QY | 1557 | gatgatgacaacactctcctctcttattcccaaatctgctcagttgtgcacaagaagcatccttaaa    | 1616 |
| Db | 1564 | gatgatgacaacactctcctctcttattcccaaatctgctcagttgtgcacaagaagcatccttaaa    | 1623 |
| QY | 1617 | acttggttatcaattacattgctgtcgtgtaagagaggaagacttggagactatgctcccttaagtcctc | 1676 |
| Db | 1624 | acttggttatcaattacattgctgtcgtgtaagagaggaagacttggagactatgctcccttaagtcctc | 1683 |
| QY | 1677 | ggccccgattgacaaggaagtataaaagcgaatatttggatacaatggtccctccagcggaattggt    | 1736 |
| Db | 1684 | ggccccgattgacaaggaagtataaaagcgaatatttggatacaatggtccctccagcggaattggt    | 1743 |
| QY | 1737 | agggaattacaanaaagttccgatttattgtgcatacaacagatgaanaacctttaaagctcgtgaa    | 1796 |

|    |      |  |      |
|----|------|--|------|
| Db | 1744 | aggaagtacaaaaagcccgacttllatcgtacatacagaatgaaaccccttaagactcgtgta      | 1803 |
| Qy | 1797 | gctattccgcattgatcatcaggaatcctctgggaaccttacttllatbvggagaagtctggagaca  | 1856 |
| Db | 1804 | gctatcagcatctgatacgaagaaactctgggaaccttacttllatbvggagaagtctggagaca    | 1863 |
| Qy | 1857 | ctgttgtatlatatllaagaatcaagcaagcagacaataaacatctccctcaggaatc           | 1916 |
| Db | 1864 | ctgttgtatlatatllaagaatcaagcaagcagacaataaacatctccctcaggaatc           | 1923 |
| Qy | 1917 | actgattgcctgtccctttgtatctcaaggaatcttaccaaaaggtgtgaaaaactttgaagat     | 1976 |
| Db | 1924 | actgattgcctgtccctttgtatctcaaggaatcttaccaaaaggtgtgaaaaactttgaagat     | 1983 |
| Qy | 1977 | ttccaattctgcgcagaggaagaatcatctcaataataatgaaacagctgctgagaagattgg      | 2036 |
| Db | 1984 | ttccaattctgcgcagaggaagaatcatctcaataataatgaaacagctgctgagaagattgg      | 2043 |
| Qy | 2037 | ccaactaatatcagatctccctggtgcctgtaccgcgtattactctagtcttgtaatactgag      | 2096 |
| Db | 2044 | ccaactaatatcagatctccctggtgcctgtaccgcgtattactctagtcttgtaatactgag      | 2103 |
| Qy | 2097 | agagatcttagcttcaagaccatcttgccctctcccatctgctctacaagaatctgtaagt        | 2156 |
| Db | 2104 | agagatcttagcttcaagaccatcttgccctctcccatctgctctacaagaatctgtaagt        | 2163 |
| Qy | 2157 | caaaagagaaacccagaaatgattctcagacagaggaatgtcatccgtttcttgatcttgat       | 2216 |
| Db | 2164 | caaaagagaaacccagaaatgattctcagacagaggaatgtcatccgtttcttgatcttgat       | 2223 |
| Qy | 2217 | gagaaacgaagctgtgtactccatacagaagaatatacaacgcttctctcccaatccagctgtga    | 2276 |
| Db | 2224 | gagaaacgaagctgtgtactccatacagaagaatatacaacgcttctctcccaatccagctgtga    | 2283 |
| Qy | 2277 | gtgcagcctctgaggaatccagaggttcccaagcttcccaatcatcagacagatccaatgtgctat   | 2336 |
| Db | 2284 | gtgcagcctctgaggaatccagaggttcccaagcttcccaatcatcagacagatccaatgtgctat   | 2343 |
| Qy | 2337 | gtttttgtatgtttgcaggttgcagtttctgttctgtgaatggatggatcgtgtatcttca        | 2396 |
| Db | 2344 | gtttttgtatgtttgcaggttgcagtttctgttctgtgaatggatggatcgtgtatcttca        | 2403 |
| Qy | 2397 | agcatctgagcacagactgtgactctcttctgtctctctctctctctgtgataataccttcaaac    | 2456 |
| Db | 2404 | agcatctgagcacagactgtgactctcttctgtctctctctctctctgtgataataccttcaaac    | 2463 |
| Qy | 2457 | aaaatggtctctatgaaagcacaaactccaccctattcccatctcagaagaactgtgtcttcatg    | 2516 |
| Db | 2464 | aaaatggtctctatgaaagcacaaactccaccctattcccatctcagaagaactgtgtcttcatg    | 2523 |
| Qy | 2517 | tcgaatggaaaaaaccagaattctcagtattctcgggggtgcacaaactcagaactcttcggaacaga | 2576 |
| Db | 2524 | tcgaatggaaaaaaccagaattctcagtattctcgggggtgcacaaactcagaactcttcggaacaga | 2583 |
| Qy | 2577 | ggcatgacccgctcttactcgaaggtttcttaattcttgacaagaacactgtgtgattatcagag    | 2636 |
| Db | 2584 | ggcatgacccgctcttactcgaaggtttcttaattcttgacaagaacactgtgtgattatcagag    | 2643 |
| Qy | 2637 | gacagttatgaaagatattctcagacatcttgcctgtgagtaaaaaaatgtccattgaaaccaaga   | 2696 |
| Db | 2644 | gacagttatgaaagatattctcagacatcttgcctgtgagtaaaaaaatgtccattgaaaccaaga   | 2703 |
| Qy | 2697 | agcttctccccaagaattcaagacaccccttagcactatgacaaaagcaattaatgtccacccca    | 2756 |
| Db | 2704 | agcttctccccaagaattcaagacaccccttagcactatgacaaaagcaattaatgtccacccca    | 2709 |
| Qy | 2757 | ccaagtctgaaacgcaccaacgaggaataactcgttaactcttccagtaagaatcaagag         | 2816 |
| Db | 2710 | -----ggaataactcgttactactctctccgttcaagatcaagag                        | 2745 |
| Qy | 2817 | gaaattgacatatgtatgataccatatcgaattgaaatgaaagaaagaatttgacattat         | 2876 |

Dh 2746 gaattgactatgatagtataccatactcagttgaaatgaagaagaatttgacattat 2805  
Qy 2877 gatgagatgaaatcagagcccccagctttcaaaaagaacaacacactattatc 2936  
Dh 2806 gatgagatgaaatcagagcccccagctttcaaaaagaacaacacactattatc 2865  
Qy 2937 gctcagatgagagcctcgatatactggaatgagtagctcccaactgttctaagaac 2996  
Dh 2866 gctgcaatgagagcctcgatatactggaatgagtagctcccaactgttctaagaac 2925  
Qy 2997 agggctcagaagtggcagtgctccctcagttcaagaagaattgtttccagaattactgat 3056  
Dh 2926 agggctcagaagtggcagtgctccctcagttcaagaagaattgtttccagaattactgat 2985  
Qy 3057 ggcctcttactcagcccttataccgttggagaactaaatgaactttgggaactctggg 3116  
Dh 2986 ggcctcttactcagcccttataccgttggagaactaaatgaactttgggaactctggg 3045  
Qy 3117 ccatataaagagcagaagtgtgaatacatactgttaactttcagaatacagccctc 3176  
Dh 3046 ccatataaagagcagaagtgtgaatacatactgttaactttcagaatacagccctc 3105  
Qy 3177 cgtccctactcctctacttctagccttattcttatgagaagatcagaagcaagagca 3236  
Dh 3106 cgtccctactcctctacttctagccttattcttatgagaagatcagaagcaagagca 3165  
Qy 3237 gaacctgaaaaaactttgtcagccttaatgaaaccaaacttacttttggaaagtgc 3296  
Dh 3166 gaacctgaaaaaactttgtcagccttaatgaaaccaaacttacttttggaaagtgc 3225  
Qy 3297 catcatactgcaaccacataaagatgagtttgaactgcaaacgctggcttattctcgat 3356  
Dh 3226 catcatactgcaaccacataaagatgagtttgaactgcaaacgctggcttattctcgat 3285  
Qy 3357 gttgaactgtaaaaaagtgtgcactcagcgcctgatttggacccctctgctcgcacact 3416  
Dh 3286 gttgaactgtaaaaaagtgtgcactcagcgcctgatttggacccctctgctcgcacact 3345  
Qy 3417 aacacatgtaaaccttgcttcatctggagaacaagtgcacatatacagaatttgcctgttttc 3476  
Dh 3346 aacacatgtaaaccttgcttcatctggagaacaagtgcacatatacagaatttgcctgttttc 3405  
Qy 3477 accactttgatgagaccaaagcttgtaactcactgtaaaatagtaagaactgcag 3536  
Dh 3406 accactttgatgagaccaaagcttgtaactcactgtaaaatagtaagaactgcag 3445  
Qy 3537 gctccctgcataatccagatggaagatcccaactttaagaagaattatcgcttccatgca 3596  
Dh 3466 gctccctgcataatccagatggaagatcccaactttaagaagaattatcgcttccatgca 3525  
Qy 3597 atcaatgctacataatgatactactacccgttgtaatagtgtcagatcaagaagatc 3656  
Dh 3526 atcaatgctacataatgatactactacccgttgtaatagtgtcagatcaagaagatc 3585  
Qy 3657 cgaatgctatctgctcagcagtgaggcaatgtaaaacatctacttacttcaagtgga 3716  
Dh 3586 cgaatgctatctgctcagcagtgaggcaatgtaaaacatctacttacttcaagtgga 3645  
Qy 3717 catgtgtcactgtacgaaaaaagaagatataaaatggcactgtatactctatcca 3776  
Dh 3646 catgtgtcactgtacgaaaaaagaagatataaaatggcactgtatactctatcca 3705  
Qy 3777 ggggttttgaagcagtggaatgttaccatccaagaacttggaatttggcggtggaatgc 3836  
Dh 3706 ggggttttgaagcagtggaatgttaccatccaagaacttggaatttggcggtggaatgc 3765  
Qy 3837 ctatattgagcagcatcactatgtctgagatgagcacttttctggtgtcaagaataag 3896  
Dh 3766 ctatattgagcagcatcactatgtctgagatgagcacttttctggtgtcaagaataag 3825  
Qy 3897 tgcacagcccccgtggaatggtcttgcacacattagagaatttccagatcaagcttca 3956  
Dh 3826 tgcacagcccccgtggaatggtcttgcacacattagagaatttccagatcaagcttca 3885

Qy 3957 ggcacatatgagcagttggccccaaagctggccagacttcaattatctccgatacataat 4016  
Dh 3886 ggcacatatgagcagttggccccaaagctggccagacttcaattatctccgatacataat 3945  
Qy 4017 gctggagaccaaaggagcccttcttctggaatcgaagtgtgactgttggcccaatgat 4076  
Dh 3946 gctggagaccaaaggagcccttcttctggaatcgaagtgtgactgttggcccaatgat 4005  
Qy 4077 attcagcgcatcaagaccaggggtgcgcgtcagaagatcttccagcctctacatctcag 4136  
Dh 4006 attcagcgcatcaagaccaggggtgcgcgtcagaagatcttccagcctctacatctcag 4065  
Qy 4137 ttatcatcatgtatagtcttctgagtggaagaagtggcagacttactcagaagaattccact 4196  
Dh 4066 ttatcatcatgtatagtcttctgagtggaagaagtggcagacttactcagaagaattccact 4125  
Qy 4197 ggaaccttaatgctctctcttgcgaatgtgatactcactggaataaaacacataatttt 4256  
Dh 4126 ggaaccttaatgctctctcttgcgaatgtgatactcactggaataaaacacataatttt 4185  
Qy 4257 aacctccaattatgtctcgatatacccglttgcacccaactcaattatagcatcgcagc 4316  
Dh 4186 aacctccaattatgtctcgatatacccglttgcacccaactcaattatagcatcgcagc 4245  
Qy 4317 actcttcgagtgaggttggatggctgtgatttaaatgattgcaatgcacttgggaatg 4376  
Dh 4246 actcttcgagtgaggttggatggctgtgatttaaatgattgcaatgcacttgggaatg 4305  
Qy 4377 gagagtaagaacatatcagatgacacagattactcttcatccttcaacaataatgttt 4436  
Dh 4306 gagagtaagaacatatcagatgacacagattactcttcatccttcaacaataatgttt 4365  
Qy 4437 gccacctgctccttcaaaaagctcgacttcaacctccaaggaagatgacttgcagaga 4496  
Dh 4366 gccacctgctccttcaaaaagctcgacttcaacctccaaggaagatgacttgcagaga 4425  
Qy 4497 cctcaggttgaaataatccaaaagtggtgtcgaagtgaacttccagaagacatagaagctc 4556  
Dh 4426 cctcaggttgaaataatccaaaagtggtgtcgaagtgaacttccagaagacatagaagctc 4485  
Qy 4557 acaagaaataactacagagtgataaaatctctgtcttccagacatgatagtgaagaagctc 4616  
Dh 4486 acaagaaataactacagagtgataaaatctctgtcttccagacatgatagtgaagaagctc 4545  
Qy 4617 ctcaatccagcagtcagaatgagcatcagtggaactctcttttccagaatgcaaaatga 4676  
Dh 4546 ctcaatccagcagtcagaatgagcatcagtggaactctcttttccagaatgcaaaatga 4605  
Qy 4677 aaggttttgaaggaaatacgaactccttcaacactgttgatgaacttcttagaccacag 4736  
Dh 4606 aaggttttgaaggaaatacgaactccttcaacactgttgatgaacttcttagaccacag 4665  
Qy 4737 ttactgactcgtactccttcaaatccaccacaggttgggtgcacagatgaccttag 4796  
Dh 4666 ttactgactcgtactccttcaaatccaccacaggttgggtgcacacagatgaccttag 4725  
Qy 4797 atggaagttctgggtgcagagcagacacttactgactgcagagttcttctgag 4856  
Dh 4726 atggaagttctgggtgcagagcagacacttactgactgcagagttcttctgag 4770  
Qy 4857 ggaatcgcaataaagaacgaataaagaacgcaaggtgtgtgggtgcgttttgcgatcca 4916  
Dh 4771 -----aataaagaatcagaagctctagagatctgtgtgtgttcttctgtgtgcg 4819  
Qy 4917 gatctagaaccccctatgtatgtgtgccaactccctctctgagcctcgtcgtccac 4976  
Dh 4820 gccgcagaagaacccctatgtatgtgtgccaactccctctctgagcctcgtcgtccac 4879  
Qy 4977 tgaagccgcccggcgaagcccggtgcgtgcgaacttgggtgcgcccgcctcgaatgag 5036  
Dh 4880 tgaagccgcccggcgaagcccggtgcgtgcgaacttgggtgcgcccgcctcgaatgag 4939

|    |      |   |      |
|----|------|---|------|
| OY | 5037 | cgcgcgagcgcgagagagagagtgcgcaacccccccccccccccctcgacgccaagt         | 5096 |
| Db | 4940 | cgagcgagcgcgag-----   | 4954 |
| OY | 5097 | gattaatgaatcgcgcaacgcgcgggagagagcggtttgcgtatctggcgcttcgcg         | 5156 |
| Db | 4955 | -----   | 4954 |
| OY | 5157 | tctctgcctcaactgcgtcgcgtcgtcgttcgcgtcgcgcgcgagatcatagctca          | 5216 |
| Db | 4955 | -----   | 4954 |
| OY | 5217 | cctcaagcgcggtatcatcgttaccacagaatacagggagataaagcagagaagaacatg      | 5276 |
| Db | 4955 | -----   | 4972 |
| OY | 5277 | agcaaaagccagcaaaagccagggacgcgtaaaagcgcggttgcgttcgatttcga          | 5336 |
| Db | 4973 | agcaaaagccagcaaaagccagggacgcgtaaaagcgcggttgcgttcgatttcga          | 5032 |
| OY | 5337 | taggcctcgccccctcgcgcgcgacatcaacaaaatcgacgctcaagltcagagtgcgaa      | 5396 |
| Db | 5033 | taggcctcgccccctcgcgcgcgacatcaacaaaatcgacgctcaagltcagagtgcgaa      | 5092 |
| OY | 5397 | ccccagagagactataagataccagcggtttcccccctggagatccctctgtgccttc        | 5456 |
| Db | 5093 | ccccagagagactataagataccagcggtttcccccctggagatccctctgtgccttc        | 5152 |
| OY | 5457 | tgcttcgcacctcgcgtctacccgataccgcttcgcgttccttccttcgcggaaagctgc      | 5516 |
| Db | 5153 | tgcttcgcacctcgcgtctacccgataccgcttcgcgttccttccttcgcggaaagctgc      | 5212 |
| OY | 5517 | gctttccaaatgcctcgcgtctgtagtctcagttcgcgttgtagtctgttcgtccaaagt      | 5576 |
| Db | 5213 | gctttccaaatgcgtcgcgtctgtagtctcagttcgcgttgtagtctgttcgtccaaagt      | 5272 |
| OY | 5577 | ggcgctgtgtgcacgaaccccccggttcaagccgcgacgcgtcgcgtctatccgtaactcg     | 5636 |
| Db | 5273 | ggcgctgtgtgcacgaaccccccggttcaagccgcgacgcgtcgcgtctatccgtaactcg     | 5332 |
| OY | 5637 | tcttgagctcaacccggttaagacagactatctgcacctgcgcgcgcgcgcgcgcgttaacag   | 5696 |
| Db | 5333 | tcttgagctcaacccggttaagacagactatctgcacctgcgcgcgcgcgcgcgcgttaacag   | 5332 |
| OY | 5697 | gattagcagagcgaggtatctagagcggtgcgtgcatacagattctcgaagtggtgcctaacta  | 5756 |
| Db | 5393 | gattagcagagcgaggtatctagagcggtgcgtgcatacagattctcgaagtggtgcctaacta  | 5452 |
| OY | 5757 | cggtctacactagaaggaacagattttgtatctcgcgtcctgcgtcgaagccagttaccttcgg  | 5816 |
| Db | 5453 | cggtctacactagaaggaacagattttgtatctcgcgtcctgcgtcgaagccagttaccttcgg  | 5512 |
| OY | 5817 | aaaagaagttgtagctcttgatcttcgaccggcaacaaacacacgcgtctgtagcggtttttt   | 5876 |
| Db | 5513 | aaaagaagttgtagctcttgatcttcgaccggcaacaaacacacgcgtctgtagcggtttttt   | 5572 |
| OY | 5877 | tgtttgcgaagcgcagattatcgcgcgcgaaaaaaaggatctccaagaagaatccttgcattct  | 5936 |
| Db | 5573 | tgtttgcgaagcgcagattatcgcgcgcgaaaaaaaggatctccaagaagaatccttgcattct  | 5632 |
| OY | 5937 | tctctacggggtctgcagctcagttgcgaacgaanaactcagttlaagggattttgtgcattgag | 5996 |
| Db | 5633 | tctctacggggtctgcagctcagttgcgaacgaanaactcagttlaagggattttgtgcattgag | 5692 |
| OY | 5997 | attatcaaaaaaggatcttcaacctagatccttttaaatataaataatgaagttttaatcaat   | 6056 |
| Db | 5693 | attatcaaaaaaggatcttcaacctagatccttttaaatataaataatgaagttttaatcaat   | 5752 |
| OY | 6057 | cctaagtatattgagttaaacttggtgcgtgcagtttccaatgtttaaagtgtagagcacc     | 6116 |
| Db | 5753 | cctaagtatattgagttaaacttggtgcgtgcagtttccaatgtttaaagtgtagagcacc     | 5812 |
| OY | 6117 | tattctcagatctgtctatttgcgttcatccatagttgcgttcagctcccgctcgtgtgagat   | 6176 |

|    |      |   |      |
|----|------|---|------|
| Db | 5813 | tatcccaagcagctcgcgtcattctcgcttcaaccataagttgcccgtacatcccgctcgtgtatagat | 5872 |
| Qy | 6177 | aactacgatacagggagggccttacacatctggccccagtgctgcgaatgataccgcgagaccc      | 6236 |
| Db | 5873 | aactacgatacagggagggccttacacatctggccccagtgctgcgaatgataccgcgagaccc      | 5932 |
| Qy | 6237 | acgcctcaacggcctccagattatcaagcaataaacccagccgaggaagggccgagcgacg         | 6256 |
| Db | 5933 | acgcctcaacggcctccagattatcaagcaataaacccagccgaggaagggccgagcgacg         | 5992 |
| Qy | 6297 | aagtggtccctcgaactcttatccgcctccacacagcttatataatgtgttcgccggagaactag     | 6356 |
| Db | 5993 | aagtggtccctcgaactcttatccgcctccacacagcttatataatgtgttcgccggagaactag     | 6052 |
| Qy | 6357 | agtaagtagttcgcgcagcttaataagtttgcgcgaacggtttgtgcacatgctacacgacatcgt    | 6416 |
| Db | 6053 | agtaagtagttcgcgcagcttaataagtttgcgcgaacggtttgtgcacatgctacacgacatcgt    | 6112 |
| Qy | 6417 | gggtgtcacgcgtcgtctgttgatggtcttcacatcagctcccggttcccaacgatacaagcg       | 6476 |
| Db | 6113 | gggtgtcacgcgtcgtctgttgatggtcttcacatcagctcccggttcccaacgatacaagcg       | 6172 |
| Qy | 6477 | agttacagtgatcccccattgtgtgcaaaaaaggggttaactctccttcgggtccctcagatcgt     | 6536 |
| Db | 6173 | agttacatgatacccccattgtgtgcaaaaaaggggttaactctccttcgggtccctcagatcgt     | 6232 |
| Qy | 6537 | tgtaagaagaatgaagtttggccgcagctgttatataacacagtggtatgagcagcactgtcaaatc   | 6596 |
| Db | 6233 | tgtaagaagaatgaagtttggccgcagctgttatataacacagtggtatgagcagcactgtcaaatc   | 6292 |
| Qy | 6597 | tctactctgcatgycacatccgtlaaagatgctttctctgtaactgtgtgactactcaaccaagtc    | 6656 |
| Db | 6293 | tcttactctgcatgycacatccgtlaaagatgctttctctgtaactgtgtgactactcaaccaagtc   | 6352 |
| Qy | 6657 | attcttgagaatatggtatgtctgcgcgcgacccagttgtctcttgcgccgcgtcaatacagggataa  | 6716 |
| Db | 6353 | attcttgagaatatggtatgtctgcgcgcgacccagttgtctcttgcgccgcgtcaatacagggataa  | 6412 |
| Qy | 6717 | taccgcgcacacatagcaggaactttaaaagtggtcactcatcttggaataaacggtttcttcggcg   | 6776 |
| Db | 6413 | taccgcgcacacatagcaggaactttaaaagtggtcactcatcttggaataaacggtttcttcggcg   | 6472 |
| Qy | 6777 | aaaactccacaggaactcttaccgcgtgttgagatccagttcgaatgataaacccactcgtgcacc    | 6836 |
| Db | 6473 | aaaaactccacaggaactcttaccgcgtgttgagatccagttcgaatgataaacccactcgtgcacc   | 6532 |
| Qy | 6837 | caactgtattctcagcactcttttaacttccacagcggtttcttggttagcgaataaacggag       | 6896 |
| Db | 6533 | caactgtattctcagcactcttttaacttccacagcggtttcttggttagcgaataaacggag       | 6592 |
| Qy | 6897 | gcaaaatgcccgcgaataaagggataagggcgacacggaacatttaaalactatcatctct         | 6956 |
| Db | 6593 | gcaaaatgcccgcgaataaagggataagggcgacacggaacatttaaalactatcatctct         | 6652 |
| Qy | 6957 | ccctttccaatbatatgtgaagcatalcttaccaggttatgtgtccaatgagcggatatacat       | 7016 |
| Db | 6653 | ccctttccaatbatatgtgaagcatalcttaccaggttatgtgtccaatgagcggatatacat       | 6712 |
| Qy | 7017 | tgaatgtattttagaataaaataacaataatagggtgtccgcgcacatttccccgaaaagtgc       | 7076 |
| Db | 6713 | tgaatgtattttagaataaaataacaataatagggtgtccgcgcacatttccccgaaaagtgc       | 6772 |
| Qy | 7077 | accctgaagcttaagaaacccatttttccatgacatataacctaataaaataggcgtataac        | 7136 |
| Db | 6773 | accctgaagcttaagaaacccatttttccatgacatataacctaataaaataggcgtataac        | 6832 |
| Qy | 7137 | gagcgccttcgtctcgcgcgcttcgtgtga-tga-cggttgaaacccctcagacatgacgt         | 7196 |
| Db | 6833 | gagcgccttcgtctcgcgcgcttcgtgtga-tga-cggttgaaacccctcagacatgacgt         | 6892 |
| Qy | 7197 | cccggaacacgttcaacagctgtctctgtaagcgga-tgc-cgggaacagacagcccggtcgaag     | 7256 |

```

Db 6893 ccggagagacggtccacagctgtctcgttaagcggatgacccgggagcagacaagcccgctcaagg 6552
Qy 7257 cgcgtcaacggggtgttcgcygggtgttcggygcttgcgttaactatgcgcatcaagcagat 7316
Db 6953 cgcgtcagcgggtgttcgcygggtgttcggygcttgcgttaactatgcgcatcaagcagat 7012
Qy 7317 tgtactgagatgtcaccatatgcggtgtgnaataccgcacagatgcgtlaaggagaata 7376
Db 7013 tgcctcggagatgcacata----- 7032
Qy 7377 ccgcatacaggaatctgtaaacgcttaatatctgtttaaatctcgcttaaatcttctttaa 7436
Db 7033 -----aaatgtcaaacgcttaatatcttgtttaaatctcgcttaaatcttctttaa 7082
Qy 7437 atcagctcatcttlttlaaccataagcgcgnaaatcgcgcanaatcccttataatcaaaagaa 7496
Db 7083 atcagctcatcttlttlaaccataagcgcgnaaatcgcgcanaatcccttataatcaaaagaa 7142
Qy 7497 tagacgcgagataagggttgcgtgtgttcgtccagtttggacaagaagatccactataaagaac 7556
Db 7143 tagcccgagataagggttgcgtgtgttcgtccagtttggacaagaagatccactataaagaac 7202
Qy 7557 gtggacatcccaacgctcaaaaggcggaanaaacgctctatcaaggcgatgtggccactacgtga 7616
Db 7203 gtggacatcccaacgctcaaaaggcggaanaaacgctctatcaaggcgatgtggccactacgtga 7262
Qy 7617 ccacacaccttaataaatttlttggatcgcaggtgcgttaaagcactaaatcggaaacct 7676
Db 7263 ccacacaccaaatcaagtcttlttgggttcgcaggtgcgttaaagcactaaatcggaaacct 7322
Qy 7677 aaagggagcccccagatttaagctcttgacggyggaagcgcgagcaacgttgcgagaaaggaa 7736
Db 7323 aaagggagcccccagatttaagctcttgacggyggaagcgcgagcaacgttgcgagaaaggaa 7382
Qy 7737 ggggaagaagacggaaggagcgcgcgcttagggcgcttgcaagttgtagcgttcaacgtcgcg 7796
Db 7383 ggggaagaagacggaaggagcgcgcgcttagggcgcttgcaagttgtagcgttcaacgtcgcg 7442
Qy 7797 gtaaccacacacccgcgcgcttaatgtgcgcgctacaaggcgct 7842
Db 7443 gtaaccacacacccgcgcgcttaatgtgcgcgctacaaggcgct 7488

RESULT 3
US-09-470-618-14
; Sequence 14, Application US/09470618
; Patent No. 6200560
; GENERAL INFORMATION:
; APPLICANT: Couto, Linda B.
; APPLICANT: Colosi, Peter C.
; TITLE OF INVENTION: Adeno-Associated Vectors for Expression of Factor VIII
; TITLE OF INVENTION: by Target Cells
; FILE REFERENCE: Avigen-04082
; CURRENT APPLICATION NUMBER: US/09/470,618
; CURRENT FILING DATE: 1999-12-22
; EARLIER APPLICATION NUMBER: 09/364,862
; EARLIER FILING DATE: 1999-07-30
; EARLIER APPLICATION NUMBER: 60/125,974
; EARLIER FILING DATE: 1999-03-24
; EARLIER APPLICATION NUMBER: 60/104,994
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 4999
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-470-618-14

```

|    |                       |  |              |                                    |
|----|-----------------------|--|--------------|------------------------------------|
|    | Best Local Similarity | 93.7%  | Pred. No. 0; |                                    |
|    | Matches 4/24;         | Conservative   | 0;           | Mismatches 243; Indels 72; Gaps 5; |
| OY | 15                    | ctgcgcgctctgcctcctcaactgagcccgaggacaaagtgccgcgacgcccgggctt | 74           |                                    |
| Db | 16                    | ctgcgcgctctgcctcctcaactgagcccgaggacaaagtgccgcgacgcccgtt    | 75           |                                    |
| OY | 75                    | gcccgaggcgcctcagtgagcagagcgagcgccagagagggagtgycacaatccatc  | 134          |                                    |
| Db | 76                    | gtctgcgccgcctcagtgagcgagcgagcgagagggagtgycacaatccatc       | 135          |                                    |
| OY | 135                   | agggggttcttcagatctcttcttaagtaaaagtgactgaaaccttaccocgtgtc   | 194          |                                    |
| Db | 136                   | agggggttcttcagatctcttcttaagtaaaagtgactgaaaccttaccocgtgtc   | 194          |                                    |
| OY | 195                   | caaaagccttgctctgtgcagagtgcttgccttcagcacaacccccaatgtcgtggc  | 254          |                                    |
| Db | 195                   | tactgtgcctcgcctcttcccgaggtgaggggagaaacctgatataagtgcaagtgc  | 254          |                                    |
| OY | 255                   | ataaggccalcagcgcatgcygaltcctcagtgctgtgtgttcagagaggaagcaaa  | 314          |                                    |
| Db | 255                   | gtgaacgctcttcttcgaaacggttgccgcgcccgcgcaagtgacgtgcagggagt   | 314          |                                    |
| OY | 315                   | ccaaccagagccttggaatgttccacaaccaatgtcagcagtggtgttctcagaag   | 374          |                                    |
| Db | 315                   | tgtgtcttaaatatcacatgcgtccagggagatgt-----tgttcttaaatatcac   | 368          |                                    |
| OY | 375                   | aaaaagcctctccacccacagcgcttgagctcgagagc---ttgacacacatgcaat  | 432          |                                    |
| Db | 369                   | gacactgacatccacatttcttcttctccacaggtatcgatccacacatgcgaat    | 428          |                                    |
| OY | 433                   | tctccacctgtctcttctctgtgccttttgcattctgacttaagtgcaccaagaat   | 492          |                                    |
| Db | 429                   | tctccacctgtctcttctctgtgccttttgcattctgacttaagtgcaccaagaat   | 488          |                                    |
| OY | 493                   | acctgtgtgcagtcggaactgtcatgtggaactatatagcgaagtgatctcgtgtg   | 552          |                                    |
| Db | 489                   | acctgtgtgcagtcggaactgtcatgtggaactatatagcgaagtgatctcgtgtg   | 548          |                                    |
| OY | 553                   | tggagcgcaagatttcctcctcctagtagtgcgaaaatcttttccatccaacccctg  | 612          |                                    |
| Db | 549                   | tggagcgcaagatttcctcctcctcctagtagtgcgaaaatcttttccatccaaccc  | 608          |                                    |
| OY | 613                   | acaaaagaacactgttttgtaagaattccacggttccacctttccaacatgcgtcaa  | 672          |                                    |
| Db | 609                   | acaaaagaacactgttttgtaagaattccacggttccacctttccaacatgcgtcaa  | 668          |                                    |
| OY | 673                   | caacctgtgagtgctctgctagagtcctacacatccagcgctgaggttaagtacag   | 732          |                                    |
| Db | 669                   | caacctgtgagtgctctgctagagtcctacacatccagcgctgaggttaagtacag   | 728          |                                    |
| OY | 733                   | tttaacttaagaanaatggtctcccatccctgtcagctctctatgtctgtgtgctct  | 792          |                                    |
| Db | 729                   | tttaacttaagaanaatggtctcccatccctgtcagctctctatgtctgtgtgctct  | 788          |                                    |
| OY | 793                   | ggaagactctcgagggagcgtgaatatgatcatcagaccagctcaaaaggagaaaga  | 852          |                                    |
| Db | 789                   | ggaaagactcttcgagggagcgtgaatatgatcatcagaccagctcaaaaggagaa   | 848          |                                    |
| OY | 853                   | ataaagttcttcctcgtgtggaagcacaataatgtctgcagcgctcctgnaaaga    | 912          |                                    |
| Db | 849                   | ataaagttcttcctcgtgtggaagcacaataatgtctgcagcgctcctgnaaaga    | 908          |                                    |
| OY | 913                   | caatggtcctcgacccaactgtgctcttaactcctacatctcttctcatgtgagcc   | 972          |                                    |
| Db | 909                   | caatggtcctcgacccaactgtgctcttaactcctacatctcttctcatgtgagcc   | 968          |                                    |
| OY | 973                   | aagacttcaatctcagggcctcatctggaagcctctactagatgttgaagaaggag   | 1032         |                                    |
| Db | 969                   | aagacttcaatctcagggcctcatctggaagcctctactagatgttgaagaaggag   | 1028         |                                    |
| OY | 1033                  | aggaaagacacagacctgtcacaaaatttactactcttctgtgtatattgtagaag   | 1092         |                                    |

Db 1029 aggaaaagacacagaccttgcacaaatltatactcttctgctgatttgaagga 1088  
QY 1093 aaagtgtgacctcagaaaacaaagaactccttgatgcagatagaggatgtgcatcgcc 1152  
Db 1089 aaagtgtgacctcagaaaacaaagaactccttgatgcagatagaggatgtgcatcgcc 1148  
QY 1153 gggccttgccctaanaatgcacacagtcnaatgtgtatgttaaacagctcctgcaggtctga 1212  
Db 1149 gggccttgccctaanaatgcacacagtcnaatgtgtatgttaaacagctcctgcaggtctga 1208  
QY 1213 ttgatgtgcacaggaagaatcagttctatgttgatgttgaaatgggcccactctgaag 1272  
Db 1209 ttgatgtgcacaggaagaatcagttctatgttgatgttgaaatgggcccactctgaag 1268  
QY 1273 tgcactcaataatccctgcgaaggtcacacattctctgttgaggaacacgcgcagctcct 1332  
Db 1269 tgcactcaataatctcctgcgaaggtcacacattctctgttgaggaacacgcgcagctcct 1328  
QY 1333 ttggaatctgcgcaataaacttctccttactgtcnaaacactcttgatggacttggacagt 1392  
Db 1329 ttggaatctgcgcaataaacttctccttactgtcnaaacactcttgatggacttggacagt 1388  
QY 1393 ttctactgttttgcatactctcttcccaacacatgcatggtgcattgaagcttatgtccaag 1452  
Db 1389 ttctactgttttgcatactctcttcccaacacatgcatggtgcattgaagcttatgtccaag 1448  
QY 1453 taagacagctgtgccagaggaaccccacacacagatgaaaaataatgaagaagcggaagact 1512  
Db 1449 taagacagctgtgccagaggaaccccacacacagatgaaaaataatgaagaagcggaagact 1508  
QY 1513 atgatagtatccttactgaatctcctgaatgtgatgtgtcaggttctgaatgaacactctc 1572  
Db 1509 atgatagtatccttactgaatctcctgaatgtgatgtgtcaggttctgaatgaacactctc 1568  
QY 1573 ctctccttaaccacaaattgcctcagttgcgaagaagacatcccttaaacacttgggtgacattaa 1632  
Db 1569 ctctccttaaccacaaattgcctcagttgcgaagaagacatcccttaaacacttgggtgacattaa 1628  
QY 1633 ttgctgtcgtgaagaggaagactctgagactatgtctcccttactgtccctgcgcgcgaatgaacaga 1692  
Db 1629 ttgctgtcgtgaagaggaagactctgagactatgtctcccttactgtccctgcgcgcgaatgaacaga 1688  
QY 1693 gttaataaagtcacaataatttgaaacaaatgtgcctcagcggaattgtgtcgggaagtacaacaaag 1752  
Db 1689 gttaataaagtcacaataatttgaaacaaatgtgcctcagcggaattgtgtcgggaagtacaacaaag 1748  
QY 1753 tccgaattatgtgcataccacagatgaagaaccttaagaacctgtgaagaaccttccagcttgat 1812  
Db 1749 tccgaattatgtgcataccacagatgaagaaccttgaagaacctgtgaagaaccttccagcttgat 1808  
QY 1813 caggaaatcttggaaccttacttcttaatgtgggaagtgtgagacacacatgttgaatatatta 1872  
Db 1809 caggaaatcttggaaccttacttcttaatgtgggaagtgtgagacacacatgttgaatatatta 1868  
QY 1873 agaatcaagcaagcagacacataatacatctacccctcagcgaatacactgtatgcgtcctt 1932  
Db 1869 agaatcaagcaagcagacacataatacatctacccctcagcgaatacactgtatgcgtcctt 1928  
QY 1933 tgaatcaaggaagtattcccaaaagtgtgaaaaacatttgaagaatttccaatttcgcgcag 1992  
Db 1929 tgaatcaaggaagtattcccaaaagtgtgaaaaacatttgaagaatttccaatttcgcgcag 1988  
QY 1993 gagaaatatltcaaalataaatagtacagtgactgtgtagaagtgtggcacaactaaatcagatc 2052  
Db 1989 gagaaatatltcaaalataaatagtacagtgactgtgtagaagtgtggcacaactaaatcagatc 2048  
QY 2053 ctgcgtgcctgcagcccgctataactctagtctcgttaataatgtgagagagatctcagcttaag 2112  
Db 2049 ctgcgtgcctgcagcccgctataactctagtctcgttaataatgtgagagagatctcagcttaag 2108  
QY 2113 gaactatggcctctctctcaatctgctacaagaagtctgtgataataaagagagaacccaga 2172  
Db 2109 gaactatggcctctctctcaatctgctacaagaagtctgtgataataaagagagaacccaga 2168

QY 2173 taatgtcagacagaagaatgtcatccctgttctctgtatattgatagagaaccgaagctggt 2232  
Db 2169 taatgtcagacagaagaatgtcatccctgttctctgtatattgatagagaaccgaagctggt 2228  
QY 2233 acctcagagaataatcacacgcttctctcccaatccagctcggagttgcacttgaagatc 2292  
Db 2229 acctcagagaataatcacacgcttctctcccaatccagctcggagttgcacttgaagatc 2288  
QY 2293 cagagtcccaagcctccacatcatgcagatcagatcaagatggtatgttttgaatgttgc 2352  
Db 2289 cagagtcccaagcctccacatcatgcagatcagatcaagatggtatgttttgaatgttgc 2348  
QY 2353 agttgtcagtttgtttgtatgaggtgtgcatactgtgttaacttgaagcatltggaacacaga 2412  
Db 2349 agttgtcagtttgtttgtatgaggtgtgcatactgtgttaacttgaagcatltggaacacaga 2408  
QY 2413 ctgacttcccttctgtcttctctctgtatatactcttaaacacaaaatgtgtctatgaag 2472  
Db 2409 ctgacttcccttctgtcttctctctgtatatactcttaaacacaaaatgtgtctatgaag 2468  
QY 2473 acacactaaccttatcccatctcaggaagaactgttctatgtcgaatggaaaaccag 2532  
Db 2469 acacactaaccttatcccatctcaggaagaactgttctatgtcgaatggaaaaccag 2528  
QY 2533 gtctaagtatctgggtgtgcacaaactcagaacttgcgaagaagagcatgtgacgccttac 2592  
Db 2529 gtctaagtatctgggtgtgcacaaactcagaacttgcgaagaagagcatgtgacgccttac 2588  
QY 2593 tgaaggttctagttgtgtcaacaaacactgtgtatattacgaagaagacgttatgaata 2652  
Db 2589 tgaaggttctagttgtgtcaacaaacactgtgtatattacgaagaagacgttatgaata 2648  
QY 2653 ttccaagatctgtgtgtgttaaaaacaaatgtccatgtgaacaaagagcttcccagat 2712  
Db 2649 ttccaagatctgtgtgtgtgttaaaaacaaatgtccatgtgaacaaagagcttcccagat 2707  
QY 2713 caagacaccctagcatagcaaaaagcaatttaatgtccaccaccagcttcttgaacgccc 2772  
Db 2708 -----ccaccagcttcttgaacgccc 2726  
QY 2773 atcaacgggaaataaactcgttactacttctcagtcagatccaagaagaaattgacatgagt 2832  
Db 2727 atcaacgggaaataaactcgttactacttctcagtcagatccaagaagaaattgacatgagt 2786  
QY 2833 ataccatatcagttgaatgaagaagagaatttgaacattatgaatgaagatgaataac 2892  
Db 2787 ataccatatcagttgaatgaagaagagaatttgaacattatgaatgaagatgaataac 2846  
QY 2893 agagcccccgcagcttccaagaagaacacagacataatttatgtctgcagtggagaagc 2952  
Db 2847 agagcccccgcagcttccaagaagaacacagacataatttatgtctgcagtggagaagc 2906  
QY 2953 tctgtgatatgtggtatgtatgtctcccaacatgttctaagaacaaagagcttaaggtgca 3012  
Db 2907 tctgtgatatgtggtatgtatgtctcccaacatgttctaagaacaaagagcttaaggtgca 2966  
QY 3013 gttcccccagttccaagaagaattgttccaggaatttaccgaatggccttctactcagc 3072  
Db 2967 gttcccccagttccaagaagaattgttccaggaatttaccgaatggccttctactcagc 3026  
QY 3073 ccttaaccgttgaagaataatgaacatatgtggaactcctggtggccatataaagagcag 3132  
Db 3027 ccttaaccgttgaagaataatgaacatatgtggaactcctggtggccatataaagagcag 3086  
QY 3133 aagtgaagaataataatcatgttgaacttccaagaatcagagcttccgtccctatctctt 3192  
Db 3087 aagtgaagaataataatcatgttgaacttccaagaatcagagcttccgtccctatctctt 3146  
QY 3193 attcagccttatcttatgtagagaatcagaggtcagaggtcagagacacttagaanaaact 3252  
Db 3147 attcagccttatcttatgtagagaatcagaggtcagaggtcagagacacttagaanaaact 3206

|    |      |  |       |
|----|------|--|-------|
| QY | 3253 | ttgtccaaagccctaaatgaaacccaacttacttcttvgaaagtgcacaacataatgycacca      | 33112 |
| Db | 3207 | ttgtcaaaagccctaaatgaaacccaacttacttcttvgaaagtgcacaacataatgycacca      | 3266  |
| QY | 3313 | ctaaagaatgaaatttgactgtcaaaagccctgggcttattcttcgaatgttgaacctggaaaag    | 33722 |
| Db | 3267 | ctaaagaatgaaatttgactgtcaaaagccctgggcttattcttcgaatgttgaacctggaaaag    | 3326  |
| QY | 3373 | atgtgcaactccagcctgtatgtggaacccctctgtgtctgcgcacactcaacacatgaaacctg    | 34322 |
| Db | 3327 | atgtgcaactccagcctgtatgtggaacccctctgtgtctgcgcacactcaacacatgaaacctg    | 3386  |
| QY | 3433 | ctcatvggagaacaagtvgaaacgtacacgaatctgtctgttcttccacacttvtgaga          | 34922 |
| Db | 3387 | ctcatvggagaacaagtvgaaacgtacacgaatctgtctgttcttccacacttvtgaga          | 3446  |
| QY | 3493 | ccaaaagctgtgtacttcaactgaaaataatggaagaacacgacgggctccctgcgaatcc        | 35522 |
| Db | 3447 | ccaaaagctgtgtacttcaactgaaaataatggaagaacacgacgggctccctgcgaatcc        | 3506  |
| QY | 3553 | agatvgaaagatcccaactctttaaagaagaaatatactgcctccatcaatcgaatgycataaa     | 36122 |
| Db | 3507 | agatvgaaagatcccaactctttaaagaagaaatatactgcctccatcaatgycataaa          | 3566  |
| QY | 3613 | tggaatacaactacgtgctctgtgtaaatggtctccagaatcgaagaatctcgatgtaactgtcca   | 36722 |
| Db | 3567 | tggaatacaactacgtgctctgtgtaaatggtctccagaatcgaagaatctcgatgtaactgtcca   | 3626  |
| QY | 3673 | gcatvgggcagagaatgaaaacaatccattcttattcttaagtggaatgtaatgttccaactgac    | 37322 |
| Db | 3627 | gcatvgggcagagaatgaaaacaatccattcttattcttaagtggaatgtaatgttccaactgac    | 3686  |
| QY | 3733 | gaaaaaaagagagatataaaaatgacacgtfacaaactccatccacgaatgttcttgaagaacg     | 37922 |
| Db | 3687 | gaaaaaaagagagatataaaaatgacacgtfacaaactccatccacgaatgttcttgaagaacg     | 3746  |
| QY | 3793 | tggaatggttacaacatcccaagcctggaattggtgcgggtggaatgyccttatvggcgagcacc    | 38522 |
| Db | 3747 | tggaatggttacaacatcccaagcctggaattggtgcgggtggaatgyccttatvggcgagcacc    | 3806  |
| QY | 3853 | tacatgtctggatvgagacaacacttttctgtgtgtacaagaataagtctaaagctccctgtg      | 3912  |
| Db | 3807 | tacatgtctggatvgagacaacacttttctgtgtgtgtacaagaataagtctaaagctccctgtg    | 3866  |
| QY | 3913 | gaaatggtctctcgacacatatagagatttccaaattacaagaacttccagagaacaatatggaacgt | 3972  |
| Db | 3867 | gaaatggtctctcgacacatatagagatttccaaattacaagaacttccagagaacaatatggaacgt | 3926  |
| QY | 3973 | gggcccacaaagctgycacagactcattatctcggaatcaatcaatgctctgycacacaaag       | 40322 |
| Db | 3927 | gggcccacaaagctgycacagactcattatctcggaatcaatcaatgctctgycacacaaag       | 3986  |
| QY | 4033 | agccctttcttgatataagtgagatctgtgtgcacaaatgatttatacagcgacacaga          | 4092  |
| Db | 3987 | agccctttcttgatataagtgagatctgtgtgcacaaatgatttatacagcgacacaga          | 4046  |
| QY | 4093 | ccacgggtgcccgtcgaagaagtctccacagcctcaaatctccaaagttaataatgtaata        | 41522 |
| Db | 4047 | ccacgggtgcccgtcgaagaagtctccacagcctcaaatctccaaagttaataatgtaata        | 4106  |
| QY | 4153 | gtcttgatvggaaagaagtvggcagaactatcgagaagaatccacatggaaccttaatgtgtct     | 42122 |
| Db | 4107 | gtcttgatvggaaagaagtvggcagaactatcgagaagaatccacatggaaccttaatgtgtct     | 4166  |
| QY | 4213 | tctttggcaatgtagatatactcgggaataaacaacataattttaacctccaattatgtg         | 42722 |
| Db | 4167 | tctttggcaatgtagatatactcgggaataaacaacataattttaacctccaattatgtg         | 4226  |
| QY | 4273 | ctcgatataccogtttgcacccaacccattatagatctgcgcacacttgcgaatgtgagtt        | 43322 |
| Db | 4227 | ctcgatataccogtttgcacccaacccattatagatctgcgcacacttgcgaatgtgagtt        | 4286  |
| QY | 4333 | tgatggtgctgtgatttaaatagtttgcgaatgacatggaatggaatggaataaacaatat        | 43922 |

|  |      |   |      |
|--|------|---|------|
| Db   | 4287 | tgatggcgctgctatttaataagcttgccagcagccattggagatggagataaagcaaat      | 4346 |
| QY   | 4393 | cggatggcacaagattactcttccatccactacttaccaatatgtttggcaactgtgtcctt    | 4452 |
| Db   | 4347 | cagatggcacaagattactcttccatccactacttaccaatatgtttggcaactgtgtcctt    | 4406 |
| QY   | 4453 | caaaaagctcgacttaccctccaaaggagagtaatgtccctggagaacctcagttgataatc    | 4512 |
| Db   | 4407 | caaaaagctcgacttaccctccaaaggagagtaatgtccctggagaacctcagttgataatc    | 4466 |
| QY   | 4513 | caaaagatgtgcctgcagtgagcttccagaagacatgtaaaagttcacagagtaactac       | 4572 |
| Db   | 4467 | caaaagatgtgcctgcagtgagcttccagaagacatgtaaaagttcacagagtaactac       | 4526 |
| QY   | 4573 | aaggaggtaaaactctgtgttccaaagcagatgtatgtgaaaggagtttctatctccagcgtc   | 4632 |
| Db   | 4527 | aaggaggtaaaactctgtgttccaaagcagatgtatgtgaaaggagtttctatccagcagtc    | 4586 |
| QY   | 4633 | aagaatggccatcagtcgagactctcttttccagaaatggcacaagtaaaagtttccaaaggaa  | 4692 |
| Db   | 4587 | aagaatggccatcagtcgagactctcttttccagaaatggcacaagtaaaagtttccaaaggaa  | 4646 |
| QY   | 4693 | atcaaagactccttcaacacctgtgtgtgaactctctagaccacgcgtactcgtactacc      | 4752 |
| Db   | 4647 | atcaaagactccttcaacacctgtgtgtgaactctctagaccacgcgtactcgtactacc      | 4706 |
| QY   | 4753 | ttcgaattacccccccagagttgggtgccaacagattgcctccaggaatggaaagttctgggt   | 4812 |
| Db   | 4707 | ttcgaattacccccccagagttgggtgccaacagattgcctccaggaatggaaagttctgggt   | 4766 |
| QY   | 4813 | ggcagaagcacaagaccctcactcgtacgcagtcgagttcttctctgaagggaatcgcaataaaa | 4872 |
| Db   | 4767 | ggcagaagcacaagaccctcactcgtacgcagtcgagttcttctctgaagggaatcgcaataaa  | 4805 |
| QY   | 4873 | gacagaataaaaagcagcaggtgtgtgtggtcgtgttctcgatccagatctagaacctta      | 4932 |
| Db   | 4806 | gaatttatcttcatctgaatagtggtgtgtttttgtgtggtgcgcgcaggaacctta         | 4865 |
| QY   | 4933 | gtgatggagttggccacacctctctctgcgcgtcgtcgcgtcactatggagccgcgcggca     | 4992 |
| Db   | 4866 | gtgatggagttggccacacctctctctgcgcgtcgtcgcgtcactatggagccgcgcggca     | 4925 |
| QY   | 4993 | aagcccgagcgctgcggcagacttctgtgcgcgcgcctcagttgaacgagcgagcgagcgag    | 5051 |
| Db   | 4926 | aagtcgcgcgcagcgccgcggtcttggccggcgagcctcagttgagcgagcgagcgagcgag    | 4984 |
| RESULT 4   |      |   |      |
| US-09-364-662-14   |      |   |      |
| : Sequence 14, Application US/09364862                                       |      |   |      |
| : Patent No. 6221349   |      |   |      |
| : GENERAL INFORMATION:   |      |   |      |
| : APPLICANT: Couto, Linda B.   |      |   |      |
| : APPLICANT: Colosi, Peter C.  |      |   |      |
| : TITLE OF INVENTION: ADENO-ASSOCIATED VECTORS FOR EXPRESSION OF FACTOR VIII |      |   |      |
| : TITLE OF INVENTION: BY TARGET  |      |   |      |
| : TITLE OF INVENTION: BY TARGET  |      |   |      |
| : FILE REFERENCE: AVIGER-03743   |      |   |      |
| : CURRENT APPLICATION NUMBER: US/09/364,862                                  |      |   |      |
| : CURRENT FILING DATE: 1999-07-30  |      |   |      |
| : EARLIER APPLICATION NUMBER: 60/125,974                                     |      |   |      |
| : EARLIER FILING DATE: 1999-03-24  |      |   |      |
| : EARLIER APPLICATION NUMBER: 60/104,994                                     |      |   |      |
| : EARLIER FILING DATE: 1998-10-20  |      |   |      |
| : NUMBER OF SEQ ID NOS: 14   |      |   |      |
| : SOFTWARE: PatentIn Ver. 2.0  |      |   |      |
| : SEQ ID NO 14   |      |   |      |
| : LENGTH: 4999   |      |   |      |
| : TYPE: DNA  |      |   |      |
| : ORGANISM: Artificial Sequence  |      |   |      |
| : FEATURE:   |      |   |      |
| : OTHER INFORMATION: Description of Artificial Sequence: Synthetic           |      |   |      |







QY 2113 gactcattggccctctctcactctcgtacacaaagatctgtagatccaagaggaaaccaga 2172  
|||||  
Db 2109 gacctaatggccctctctcctccctcgtacacaaagatccgttagatccaagaggaaaccaga 2168  
2173 taatgtcagacaaagaggaatgtcatccgtctctcgtatcttgatctgagacagccgaagctggt 2232  
|||||  
Db 2169 taatgtcagacaaagaggaatgtcatccgtctctcgtatcttgatctgagacagccgaagctggt 2228  
2223 accccaagagaataatacaagcttctccccaatccagctcgaggtgcagcttgaagatc 2232  
|||||  
Db 2229 accccaagagaataatacaagcttctccccaatccagctcgaggtgcagcttgaagatc 2288  
2293 cagagcttcccaagcctcccaacatcagcagcatcaatggtctatgcttttgaatgttgc 2352  
|||||  
Db 2289 cagagcttcccaagcctcccaacatcagcagcatcaatggtctatgcttttgaatgttgc 2348  
2353 agttgtcagcttctgtctcgtcataggtggtcatatcgttacatctaaagcatlgtgagcacaga 2412  
|||||  
Db 2349 agttgtcagcttctgtctcgtcataggtggtcatatcgttacatctaaagcatlgtgagcacaga 2408  
2413 ctgagcttccctctgtctctctctcgtgataataccttccaacacacaaatggtctatgaag 2472  
|||||  
Db 2409 ctgagcttccctctgtctctctctcgtgataataccttccaacacacaaatggtctatgaag 2468  
2473 acacactcaaccttattcccatctcagagagaatctgtcttcatatgtctatgagaaaccag 2532  
|||||  
Db 2469 acacactcaaccttattcccatctcagagagaatctgtcttcatatgtctatgagaaaccag 2528  
2533 gctcatgtagatctggtgtgcacacactcgaacttcggaacagagcatgtgcccgttac 2592  
|||||  
Db 2529 gctcatgtagatctggtgtgcacacactcgaacttcggaacagagcatgtgcccgttac 2588  
2593 tgaaggttctcagttgtgtgacaaagacactggtgtattatcagagagacgttatgaaagata 2652  
|||||  
Db 2589 tgaaggttctcagttgtgtgacaaagacactggtgtattatcagagagacgttatgaaagata 2648  
2653 tttaagacatactgtctgtagtaaaacaatgcatctgaacccaagaagcttctcccaagatt 2712  
|||||  
Db 2649 tttaagacatactgtctgtagtaaaacaatgcatctgaacccaagaagcttctcccaagatt- 2707  
2713 caaagaccctcagacactagcgaaagaatttaattgcaaccccccacgcttctgaaagccc 2772  
|||||  
Db 2708 -----caacacagcttctgaaagccc 2726  
2773 atcaacgggaaataacccgtactactctctcagtcagatcaaagagaaatgtgactatgt 2832  
|||||  
Db 2727 atcaacgggaaataacccgtactactctctcagtcagatcaaagagaaatltgactatgt 2786  
2833 ataccatacagttgaaatgaagaagaagatttgaacatttgaatgagatgaaatc 2892  
|||||  
Db 2787 ataccatacagttgaaatgaagaagaagatttgaacatttgaatgagatgaaatc 2846  
2893 agagcccccgagcttccaagaagaaacaacgacacatttattgtctgcagttgaaagggc 2952  
|||||  
Db 2847 agagcccccgagcttccaagaagaaacaacgacacatttattgtctgcagttgaaagggc 2966  
2953 tctgtgattatgtggaatgagtagctccccaacatgttctaagaacaagggtccaaggtgca 3012  
|||||  
Db 2907 tctgtgattatgtggaatgagtagctccccaacatgttctaagaacaagggtccaaggtgca 2966  
3013 gtgtccctcagttcaagaagaagtttctcagagaaattacgaggtgctcttactaagc 3072  
|||||  
Db 2967 gtgtccctcagttcaagaagaagtttctcagagaaattacgaggtgctcttactaagc 3026  
3073 ccttataccgtggaagaactaaatgaacatttggagcctctggtggcatatataagagcag 3132  
|||||  
Db 3027 ccttataccgtggaagaactaaatgaacatttggagcctctggtggcatatataagagcag 3086  
3133 aagttgaagaataatacatgtgtaacttccagaaatcagggcctctgctccatctctct 3192  
|||||  
Db 3087 aagttgaagaataatacatgtgtaacttccagaaatcagggcctctgctccatctctct 3146

QY 3193 attcagccttatttcttattatgaggaaatcagaggcgaaggacgaacactagaanaaact 3252  
|||||  
Db 3147 attcagccttatttcttattatgaggaaatcagaggcgaaggacgaacactagaanaaact 3206  
3253 ttgtcaagcctaataatgaacccaacttacttcttgaagagtgcacaatcatatgtgcacca 3312  
|||||  
Db 3207 ttgtcaagcctaataatgaacccaacttacttcttgaagagtgcacaatcatatgtgcacca 3266  
3313 ctaaaagtgaattgtgactgcgaagcctggtctatttctctgtatggttgaacctggaanaag 3372  
|||||  
Db 3267 ctaaaagtgaattgtgactgcgaagcctggtctatttctctgtatggttgaacctggaanaag 3326  
3373 atgtgcactcagggcctgatttgaacccctctgtgtctgcacacataacacacttgaacctg 3422  
|||||  
Db 3327 atgtgcactcagggcctgatttgaacccctctgtgtctgcacacataacacacttgaacctg 3386  
3433 ctcatgtggaagaagaatgagcaatgtgcctggttcttccacatcttattatga 3492  
|||||  
Db 3387 ctcatgtggaagaagaatgagcaatgtgcctggttcttccacatcttattatga 3446  
3493 ccaaaagctgtgtacttcaactgcgaanaataatggaagaagaactgcagggctccctgcataatcc 3552  
|||||  
Db 3447 ccaaaagctgtgtacttcaactgcgaanaataatggaagaagaactgcagggctccctgcataatcc 3506  
3553 agatggaagaatccaccttttaagaagaattatgccttccatgcacataatgtgctacataa 3612  
|||||  
Db 3507 agatggaagaatccaccttttaagaagaattatgccttccatgcacataatgtgctacataa 3566  
3613 tgaatacactacactcgtctgtagtaatgtctcagagatccaagagatcgaatgtatctgtca 3672  
|||||  
Db 3567 tgaatacactacactcgtctgtagtaatgtctcagagatccaagagatcgaatgtatctgtca 3626  
3673 gcatgtggcagcaatgaaacaatccatctcatatcattcagttgagcatgtgttcaactgttac 3732  
|||||  
Db 3627 gcatgtggcagcaatgaaacaatccatctcatatcattcagttgagcatgtgttcaactgttac 3686  
3733 gaaaaaaagaagagtataaaatgtgcactgtgacaactctctccaggtgttcttgaagcag 3792  
|||||  
Db 3687 gaaaaaaagaagagtataaaatgtgcactgtgacaactctctccaggtgttcttgaagcag 3746  
3793 tggaaatgttaccatccaagaacttggaaatttggcgggtggaagtgttcaatgtgcagatcc 3852  
|||||  
Db 3747 tggaaatgttaccatccaagaacttggaaatttggcgggtggaagtgttcaatgtgcagatcc 3806  
3853 tactgtctggagatgagacacatttctgtgtgtacagcaataaagtgtcagagctccctgtg 3912  
|||||  
Db 3807 tactgtctggagatgagacacatttctgtgtgtacagcaataaagtgtcagagctccctgtg 3866  
3913 gaatggctctcggacacattagagatttccagattacagcttccagggacaatatgagaggt 3972  
|||||  
Db 3867 gaatggctctcggacacattagagatttccagatttccagattacagcttccagggacaatatgagaggt 3926  
3973 gggccccaagcttggccaagacttcatattatccggatcaatcaatcgtctggaagccaag 4032  
|||||  
Db 3927 gggccccaagcttggccaagacttcatattatccggatcaatcaatcgtctggaagccaag 3986  
4033 agcccttctcttggatcaaggttgaatcgtgtggtgcaccaatgatatccaggtcacaaga 4092  
|||||  
Db 3987 agcccttctcttggatcaaggttgaatcgtgtggtgcaccaatgatatccaggtcacaaga 4046  
4093 cccaagggtgcgcgtcagaagtttccagagctctcagatctccatctccagttatcatatgata 4152  
|||||  
Db 4047 cccaagggtgcgcgtcagaagtttccagagctctcagatctccatctccagttatcatatgata 4106  
4153 gtctctgtggaagaagaatggaagacttatcgaggaatctccactgcgaaccttaatgtctc 4212  
|||||  
Db 4107 gtctctgtggaagaagaatggaagacttatcgaggaatctccactgcgaaccttaatgtctc 4166  
4213 tcttggcaatgtgagatcatctgtggaataaacaacaatattttaacctccaatattgt 4272  
|||||  
Db 4167 tcttggcaatgtgagatcatctgtggaataaacaacaatattttaacctccaatattgt 4226  
4273 ctcgatacacatccgttctgacccaactcatattatagcatctgcagcactcttcgcatgtgagc 4332



|||||  
Db 541 GTGGACCTGGTAAAGACTGAATTTCAGGCTCATTTGGACCCCTACTACTATGTAGAGAA 600  
QY 1020 gggagctctggccaaggaagaacacagaccttgcacaattatatactactcttctgcgtta 1079  
Db 601 GGGAGCTTGCCCAAGGAAAGACACAGACCTTGCACAAATTTATACACTTTTTCGCTGA 660  
QY 1080 ttgatgaaggaaagtctggacctcagaaacaaagaaaccttgatgcaggaataggaat 1139  
Db 661 TTTGATGAAGGAAAGTGTGGCACTCGAAGACAAAGACTCTTGATGCGAGATAGGGAT 720  
QY 1140 gctgcactctcgggccttggcctaaatgacacagltcaatggttaagttaacagagctc 1199  
Db 721 GCTGATCTGCTGGGCGCTGGCCCTAAATGCACACACTCAATGATTATGTAACAGAGTCT 780  
QY 1200 ctgcagagctctgattggaatgcacaggaataatcagttcatttggaatgattggaatggc 1259  
Db 781 CTGCGCAGCTGTGATGTGATCCACAGGAAATCAGTCTATTGGCATGTGATGGAAATGGGC 840  
QY 1260 accactctgaagtgcactcaatattccctcgaaggtcacacattctctgtgaggaacct 1319  
Db 841 ACCACTCTCTAAGTGCACCTAATATTTCTGGAAGTCCACACTTTCTTGAGAGAACCAT 900  
QY 1320 cgcgaagcgtctcttggaaatctgcaccaataacttcccttactgtctcaaacactcttgatg 1379  
Db 901 CGCAGGCGTCTTGGAATCTGCGCAATCTGCTCTCTACTGCTCAACACTCTTGATG 960  
QY 1380 gaacttggagaagttctactcgttttgcatactctccccaacaatgattggaatgaa 1439  
Db 961 GACCTTGGAGAGTTTCTACAGTTTGTGCATATCTCTCCACCAACATGTATGGCATGGAA 1020  
QY 1440 gcttatgtcaaaatagacagcgtctccagaagaaaccccaactacgaatgaaataatgaa 1499  
Db 1021 GCTTATGTCAAAATAGACACTGTCCAGAGAACCCCACTAGCAATGAAAAATATGAA 1080  
QY 1500 gaagcggaaactatgattgattcttactgattctgaaatggaatggtgcaggttgcatt 1559  
Db 1081 GAAGCGGAAACATATGATGATCTTACTGATTCGAAATGATGATGGATGGATGTTGAT 1140  
QY 1560 gatgaacactctctctcttacttccaaatcgctcagtttgcagaagaactcctaaact 1619  
Db 1141 GATGACAACTCTCTCTCTTATTCCAAATTCGCTCAGTTCCCAAGAAAGCTCTTAAACT 1200  
QY 1620 tgggtacattacattgctgtctgaagaaggaagacttggagactatgctcccttaagcctccgc 1679  
Db 1201 TGGGTATATTACATTGCTGCTGGAAGAGAGACTGGAGACTATGCTCCCTTAGTCTCGCC 1260  
QY 1680 cccgatgacagaagtataaaatgcaatttgaacaatgccccctcagcggatgtgtagg 1739  
Db 1261 CCGGATGACAGAGTTATATAAGTCAATATTGAAACATGCGCCCTCAGCGGATTTGTAAG 1320  
QY 1740 aagtacaaaaaagctcgaatttatggtacatacagaatgaaaccttaagactcgtgaagct 1799  
Db 1321 AAGTACAAAAAATGCGATTATGCGATACACAGATGAAACCTTTAAGACTCTGTGAAGCT 1380  
QY 1800 attcagatgaactcagaatcttggagaccttacttattgagggaagtggagacacactg 1859  
Db 1381 ATTACAGATGAATCAGGAACCTTGGGACCTTTACTTTATGGGGAAGTTGGAGACACACTG 1440  
QY 1860 ttgaattatattaagaatcaagaacagacacataaacaattacacctcaagaaactact 1919  
Db 1441 TTTGATTATATTAAAGATCAAGACAGACATATTAACATCTACCCCTCAGGAATACACT 1500  
QY 1920 gatctgcgtcctcttgattatcaagaagatcaaaaaggtgtaaaacatttgaagatttc 1979  
Db 1501 GATGTCCGTCCTTTGTATTAAGGAGTTTACCAAAGGTTTAAACATTTGAAGGATTTT 1560  
QY 1980 ccaattctgcagagaaatattcaaatataaattgagacgtgacttgaagaatgagcca 2039  
Db 1561 CCAATTCTGCAGAGAAATATTCAAAATATGAATGACACTGTAGAAAGATGGGCCA 1620  
QY 2040 actaatcagaatccctcggtgcctgaccgctattactctagttcgttaatatgagaga 2099  
|||||

Db 1621 ACTAAATCAGATCCTGGTGTGACCCGCTATTACTCTAGTTTCGTTAATATGAGAGA 1680  
QY 2100 gatctgcttcaagactcatttggcctctctctcatctgtctcaaaaagaatctgtagatcaa 2159  
Db 1681 GATCTAGCTTTCAGGACTCATTTGGCCCTCTCTCATCTGCTACAAAAGATCTGTAGATCAA 1740  
QY 2160 agaggaacacagaataatgtcagaagaagaatgtcaatccctgttctctgatttgaatgag 2219  
Db 1741 AGAGAAACACAGATTAATGTACAGCAAGAGAAATGTCACTCTGTTTCTGTATTTGATGAG 1800  
QY 2220 aaccgaagcttgcactcctacagagaalatacaacgcttctccccaatccagctggagt 2279  
Db 1801 AACCGAAGCTGATACCTCACAGAAATATACAAAGCTTTCTCCCAATCCAGCTGGAGTG 1860  
QY 2280 caacttgaggatccagagttccaaagctccaacataatgcacagatcaatgtgtatgct 2339  
Db 1861 CAGCTTGAGGATCCAGAGTTCCAAAGCTCCACATCATGACAGCATCAATGGCTATGTT 1920  
QY 2340 ttctgaattgtcagttgtcagtttcttgcatagggttgcatacttggatacttctaagc 2399  
Db 1921 TTTGATATGTTGCAAGTTTGATGTTGTTTGCATGAGGTGCACTGTGATCTTCAAGC 1980  
QY 2400 attgagacagactgacttctcttctgtcttctctctctgatalataccttcaaacacaa 2459  
Db 1981 ATTGAGACACAGACTGACTCTTCTGCTCTCTCTCTGATATACCTTCAAAACACAAA 2040  
QY 2460 atggtctatgaagaacacaccacacctattcccaattctcaggaagaactgtctcagctg 2519  
Db 2041 ATGCTATGAAACACACACCTACCTATTCCTCATTCAGAGAAACATGCTTCAATGCTG 2100  
QY 2520 atgaaacccagagctatgattcttgggtgtgcacaacccagacttggagaagaagc 2579  
Db 2101 ATGGAACACCCAGCTGATGATTTCTGGGTGCCACACACTCAGACTTTGGAAACAGAGCC 2160  
QY 2580 atgacgcttactgaaagtttctagttgtgcagaagaacacgtgtgattatcaagaagac 2639  
Db 2161 ATGACCGCCTTACTGAAAGTTTCTAGTTTGACACAGAACACTGCTGATTTATCGAGGAC 2220  
QY 2640 agttatgaagatatctcagaatctgtctgagtaaaacaatccatttgaaccaagaagc 2699  
Db 2221 AGTTATGAAGATATTTGAGATCTTGTGAGTGAATAAACATGCAATGAAACAGAAAGC 2280  
QY 2700 ttctccagaatlcagaaccccttaagacactagcacaagaactttaaagccaccca 2759  
Db 2281 TTCTCCAGATTCAGAGACACCTTACGACTAGSCAAAGCAATTTAATGCCACCCACCA 2340  
QY 2760 gttctgaacgcacalcaacgggaataactcgtactactcttcaagtcagaatcaagaaga 2819  
Db 2341 GTCTTGAACGCGCATCAACGGGAATAACTCGTACTCTTCAGTCAGATCAAGAGAA 2400  
QY 2820 attgactatgattatccatataatgattgaaatgaagaaggaatttgaactttatgat 2879  
Db 2401 ATTGACTATGATATACCATATCATGTTGAATGAAGAGAAATTTGACATTTATGAT 2460  
QY 2880 gaagatgaaaatcagaagcccccagacttccaagaagaaacaagacactatattatgct 2939  
Db 2461 GAGAGTGAATAATAGAGCCCCCGACCTTTCAAAAGAAAAACGACACTTTTATTGCT 2520  
QY 2940 gaagtggagaagctctggaatttatggatgagtagctcccccacatgltcctaagaacag 2999  
Db 2521 GCAGTGGAGAGCGCTTGCGATTATGGAGTAGAGTCCCCACATGTTTCTTAAGAAACAG 2580  
QY 3000 gttcagaagtggaggttccctcagttcaagaagaagttgtttccaagaattttagatgagc 3059  
Db 2581 GCTCAAGTGGAGTGTCTCCCTGACTCAAGAAAGTTGTTTCCAGGAATTTACTGATGCG 2640  
QY 3060 tcccttaccagcccttataccgttggagaactaaatgaacatttggagactccctggagcca 3119  
Db 2641 TCTTTACTACGAGCCCTTATACGTTGAGAGAACTTAATGAACATTTGGAGACTCCCGGGCCA 2700  
QY 3120 tatataagacagaagtgtgaagataatcatatgtaacttccagaaatcaagcctctgct 3179  
Db 2701 TATATGAAGACGAAAGTTGAAGATATATCATGTTCAACTTTCAGAAATCAGGCTCTCTGCT 2760  
|||||



```
REFERENCE/DOCKET NUMBER: TTI-138
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 9354 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2965..7378
US-08-683-839B-2

Query Match          55.5%; Score 4411.6; DB 1; Length 9354;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 4420; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 415 ccaacatgcaataagagctctccactgctcttcttgagcctttgagattcgtctta 474
    |||||
Db 2960 CCACCATGGTTTATGAGCTCTCCACCCTCTCTTCTGCTGCTTTGCGATTCTGCTTTA 3019

QY 475 gtagccacgaataatactactcctggtgagcagtgagtgagtgagtgagtgagtgag 534
    |||||
Db 3020 GAGCCACCGAAGATATCTACTGGGTGAGTGAGTGTATGAGTGAGTGAGTGAGTGAGTG 3079

QY 535 atctcgtagctgctgctgtagcagcaagattctctctagagtgccaaaattcttccat 594
    |||||
Db 3080 ATCTGGTAGAGCTGGCTGTGAGACGAAGATTCTCTCTAGAGTGCCAAATCTTTCCAT 3139

QY 595 tcaaacctcagtcgtgtaacaaaagctcgttctgtagaatcagcgttcaacttca 654
    |||||
Db 3140 TCAMACCTCAGTCGTGTACAAAAGACTCTGTTGTAGATTCAAGGTTCACCTTTCA 3199

QY 655 acatgcctaacgcaagcacccttgatggtctgctgtagtcttaccatccagagcagag 714
    |||||
Db 3200 ACATGCCCTAAGCCAAAGGACCCTGATGGGTGCTGCTAGGCTCTTACCATCCAGGCTGAG 3259

QY 715 ttatgatacagtgatcattactaacttaagaacatggtctccatccctgctcagtcctatg 774
    |||||
Db 3260 TTTATGATACAGTGTGATTAACCTTAAGACATGAGTGTCCCATCTCTGACAGTCTTCAAG 3319

QY 775 cgttggtgtatctactactggaagctctcgaggagcgtgaatatgatatgcagaccagtc 834
    |||||
Db 3320 CTGTTGGTGTATCTACTGTGAAGCTTCTGAGGAGCTGAATATGATGATCAGACAGTC 3379

QY 835 aaagagagaagaagaatgaagctctccctggtgagagcagatataatgctcagag 894
    |||||
Db 3380 AAAGGAGAAAGAAAGATGAATGAAGTCTTCTCTGTGGAAGCCATGATATGTGTGCAAG 3439

QY 895 tccctgaagaagaatggtccaatgagcctctgaccacactgctcttaccataatactt 954
    |||||
Db 3440 TCCTGAAAGAGATGGTCCAAATGGCCTGTGACCACCTGTGCTTACTACTCATATCTTT 3499

QY 955 ctcaatgtagaactggttaaaagactgaaatcagggcctcaatgtaggagcctactagatgta 1014
    |||||
Db 3500 CTCATGCTGACCTGTGAAAAGACTTGAAATTCAGGCTCATTTGAGCCCTACTAGTATGTA 3559

QY 1015 gagaagaggagcttgcccaagaaagacagacacttgacaaaattataactacttttg 1074
    |||||
Db 3560 GAGAAAGGAGCTGTGGCCAAAGAAAGACAGACCTTGTGACAAATTTATATCTTTTGG 3619

QY 1075 ctgtatlttgaagaagaaagtgtggcactcagaacaagaactcctgtatgcaagata 1134
    |||||
Db 3620 CTGTATTTTATGAGGAGAAAGTGTGGCACTCAGAAACAAAGAACTCTGTATGACAGGATA 3679

QY 1135 gggatgctgcatctgctgagccttgagcctaaatgcaacagtcgaatggttatgttaaca 1194
    |||||
Db 3680 GGGATGCTCATCTGCTGGGCTGTGCCCTGAATAATGACACAGTCAATGTTATGTAAACA 3739
```

```
QY 1195 ggtctgcccaggctcgtatgtagtgccacaggaatcagtcctatggtcatgtattggaa 1254
    |||||
Db 3740 GGTCTCTGCCAGGTCTGATTGGATGTCACAGGAATACGCTATTGGCATGTATTGGAA 3799

QY 1255 tgggacacactcctggaatgtagcactcaatatctcctgaaaggttcaacatctctgttgaaga 1314
    |||||
Db 3800 TGGGACACACTCTGAAAGTGCACCTCAATATCTCTCGAAGGTCAACATTTCTTGTGAGGA 3859

QY 1315 accatgcccagggtcctcttggaaatctgcgaataacttctcctactgctcaaacctct 1374
    |||||
Db 3860 ACCATGCCACAGGGCTCTTGGAAATCTCGCAATTAATCTTCTACTGTGCTCAAAACCTCT 3919

QY 1375 tgaatgaccttgagacagtttctactgttttgtcatatctctccacacacatgatggcca 1434
    |||||
Db 3920 TGATGACCTTGGACAGTTTCTACTGTGTTTGTATATCTTCTCCACCAACATGATGGCA 3979

QY 1435 tggaaacttatgtcacaagtagacagctgtccagaagaaaccccaactagaaatgaaaaata 1494
    |||||
Db 3980 TGGAACTTATGTCAAAAGTAGACAGCTGTCCAGAGAACCCCAACTACGAATGAAAAATA 4039

QY 1495 atgaagaagcggaaagactatgatagtatcttactgtatctgaatgtagtgcaggt 1554
    |||||
Db 4040 ATGAAGAAGCGGAAGACTATGATGATGATCTTACTGAAATGGAATGTGTGAGGT 4099

QY 1555 ttgatgatgacaactctctccttcttctatccaaatltgcgtcagttgcccagaagacatcccta 1614
    |||||
Db 4100 TTGATATGACAACTCTCTCTCTTATTCAAATTTGCTCAGTTGTCAGAAACATCTTA 4159

QY 1615 aaacttgggttacaattacatgtcgtcgaagagagagcgtggaactatgctcccttatgcc 1674
    |||||
Db 4160 AAACCTTGGGTACTTATCAATTCCTGCTGGAAGAGAGAGACGTGAGCTTGTGCTTGTGTC 4219

QY 1675 tcgccccgaatgcagaagttataaagtcaaatatttgaacaatggtccctcagcagatg 1734
    |||||
Db 4220 TCGCCCCGATGACAGAAAGTTATTAAGAAAGTCAATATTTGAAACATGTGCCCCCTCACGGATTG 4279

QY 1735 gtagaagatgacaaaagaagtcggaattatgcatagacagatggaaccttaagactcgtg 1794
    |||||
Db 4280 GTAGGAAGTACAAAAAGTCCGATTTATGCGATACACAGTGAACCTTTAAGACTCGTG 4339

QY 1795 aagctatcagaatgaatcagaagatcttggaaccttacttcttgaagaggttggagaca 1854
    |||||
Db 4340 AAGCTATTGACGATGAATGAGGAATCTTGGGACTTACTTATGGGGAAGTTGGAGACA 4399

QY 1855 cactgttgatataatttaagaatcaagaagcagacacataacatctacccctcaaggaa 1914
    |||||
Db 4400 CACTGTTGATTATTAATTAAGAAATCAAGCAAGCAACCATATATACATCTCCCTCACGGGA 4459

QY 1915 tcaatgagtcgcgtccttctgtaataaggagatctaccaaaaggtgfaaaacatttgaag 1974
    |||||
Db 4460 TCACGTATGTCCTCTCTTGTATTCAGAGAGATTTACCAAAAGGTGTAACATTTGAAG 4519

QY 1975 atttccaaattcgcagagagaataatccaatataatggaagcagcgtgtagaagatg 2034
    |||||
Db 4520 ATTTTCCAAATTTCGCCAGAGAAATTTCAAAATATTAATGAGACAGTGTGAGAGATG 4579

QY 2035 ggcacaactaaatcagatccctcgtggtgcgtgaacccgcatatacttactagttctgtaataatg 2094
    |||||
Db 4580 GGCCAACTTAATCAGATTCCTCGGTGCTGACCCGCTATTACTGTATTTCTGTTAATATG 4639

QY 2095 agagaagatctagcttaagacatcattggccctctcctcaatctgctacaagaatcgtgtg 2154
    |||||
Db 4640 AGAGAGATCTAGCTTGAAGACTCATTTGGCCCTCTCTCATCTGCTTCAAAAGAAATCTGTAG 4699

QY 2155 atcaagaggaaacagaaatgtaacagacaagaagagatgtcaatccgtttcttctgtatgg 2214
    |||||
Db 4700 ATCAAGAAGAGAAACAGATATATGTGACACAAGAGAAATGTATTCGTTTGTGATTTGG 4759

QY 2215 atgagaaccgaagcgtgtatcctccacagagaataatacagcgttcttcccacatccagctg 2274
    |||||
Db 4760 ATGAGAAACCGAAGCTGTGCTACCTACAGAGAAATATACAAAGCTTTTCTCCCAATTCAGCTG 4819

QY 2275 gagtgagcttgagatccagaggttccaaagcctccaaatcaatgacagcatcaatggtct 2334
```

|   |      |     |                            |                          |                          |      |
|---|------|-----|----------------------------|--------------------------|--------------------------|------|
| D | 4820 | GA  | GTGACGCTTGAGGATCCAGAGTTC   | CAAGCCTCCAACTATCATGCA    | GCATCATATGCT             | 4879 |
| O | 2335 | AT  | TTTTTGAATTTGCAATTTGCAATTTG | CAATTTGCAATTTGCAATTTG    | CAATTTGCAATTTG           | 2394 |
| D | 4880 | AT  | TTTTTGAATTTGCAATTTGCAATTTG | CAATTTGCAATTTGCAATTTG    | CAATTTGCAATTTG           | 4939 |
| O | 2395 | TA  | AGCATTTGAGCAGCAGTACTCTCTCT | CTCTCTCTCTCTCTCTCTCTCTCT | CTCTCTCTCTCTCTCTCTCTCTCT | 2454 |
| D | 4940 | TA  | AGCATTTGAGCAGCAGTACTCTCTCT | CTCTCTCTCTCTCTCTCTCTCTCT | CTCTCTCTCTCTCTCTCTCTCTCT | 4999 |
| O | 2455 | ACA | AAATGCTATGAGACACACTACCTAT  | TCCATTCCTCAGAGAACTGCTTCA |                          | 2514 |
| D | 5000 | ACA | AAATGCTATGAGACACACTACCTAT  | TCCATTCCTCAGAGAACTGCTTCA |                          | 5059 |
| O | 2515 | TC  | TCGATGGAACCCAGGCTATGATTTCT | GGGTCCTCAGACTCTTGGAACA   |                          | 2574 |
| D | 5060 | TC  | TCGATGGAACCCAGGCTATGATTTCT | GGGTCCTCAGACTCTTGGAACA   |                          | 5119 |
| O | 2575 | GA  | AGCATTTGAGCAGTACTCTCTCTCT  | CTCTCTCTCTCTCTCTCTCTCTCT |                          | 2634 |
| D | 5120 | GA  | AGCATTTGAGCAGTACTCTCTCTCT  | CTCTCTCTCTCTCTCTCTCTCTCT |                          | 5179 |
| O | 2635 | AG  | CAATGCTTGAAGATTTTCAACATCT  | TGAGTAAACATGCTTGAACCA    |                          | 2694 |
| D | 5180 | AG  | CAATGCTTGAAGATTTTCAACATCT  | TGAGTAAACATGCTTGAACCA    |                          | 5239 |
| O | 2695 | GA  | AGCATTTGAGCAGTACTCTCTCTCT  | CTCTCTCTCTCTCTCTCTCTCTCT |                          | 2754 |
| D | 5240 | GA  | AGCATTTGAGCAGTACTCTCTCTCT  | CTCTCTCTCTCTCTCTCTCTCTCT |                          | 5299 |
| O | 2755 | CA  | AGCATTTGAGCAGTACTCTCTCTCT  | CTCTCTCTCTCTCTCTCTCTCTCT |                          | 2814 |
| D | 5300 | CA  | AGCATTTGAGCAGTACTCTCTCTCT  | CTCTCTCTCTCTCTCTCTCTCTCT |                          | 5359 |
| O | 2815 | AG  | GAATTTGAGCAGTACTCTCTCTCT   | CTCTCTCTCTCTCTCTCTCTCTCT |                          | 2874 |
| D | 5360 | AG  | GAATTTGAGCAGTACTCTCTCTCT   | CTCTCTCTCTCTCTCTCTCTCTCT |                          | 5419 |
| O | 2875 | AT  | GAATTTGAGCAGTACTCTCTCTCT   | CTCTCTCTCTCTCTCTCTCTCTCT |                          | 2934 |
| D | 5420 | AT  | GAATTTGAGCAGTACTCTCTCTCT   | CTCTCTCTCTCTCTCTCTCTCTCT |                          | 5479 |
| O | 2935 | TC  | TCGATGGAACCCAGGCTATGATTTCT | GGGTCCTCAGACTCTTGGAACA   |                          | 2994 |
| D | 5480 | TC  | TCGATGGAACCCAGGCTATGATTTCT | GGGTCCTCAGACTCTTGGAACA   |                          | 5539 |
| O | 2995 | AC  | AGCATTTGAGCAGTACTCTCTCTCT  | CTCTCTCTCTCTCTCTCTCTCTCT |                          | 3054 |
| D | 5540 | AC  | AGCATTTGAGCAGTACTCTCTCTCT  | CTCTCTCTCTCTCTCTCTCTCTCT |                          | 5599 |
| O | 3055 | AT  | GAATTTGAGCAGTACTCTCTCTCT   | CTCTCTCTCTCTCTCTCTCTCTCT |                          | 3114 |
| D | 5600 | AT  | GAATTTGAGCAGTACTCTCTCTCT   | CTCTCTCTCTCTCTCTCTCTCTCT |                          | 5659 |
| O | 3115 | GC  | CAATGCTTGAAGATTTTCAACATCT  | TGAGTAAACATGCTTGAACCA    |                          | 3174 |
| D | 5660 | GC  | CAATGCTTGAAGATTTTCAACATCT  | TGAGTAAACATGCTTGAACCA    |                          | 5719 |
| O | 3175 | TC  | TCGATGGAACCCAGGCTATGATTTCT | GGGTCCTCAGACTCTTGGAACA   |                          | 3234 |
| D | 5720 | TC  | TCGATGGAACCCAGGCTATGATTTCT | GGGTCCTCAGACTCTTGGAACA   |                          | 5779 |
| O | 3235 | CA  | AAATGCTATGAGACACACTACCTAT  | TCCATTCCTCAGAGAACTGCTTCA |                          | 3294 |
| D | 5780 | CA  | AAATGCTATGAGACACACTACCTAT  | TCCATTCCTCAGAGAACTGCTTCA |                          | 5839 |
| O | 3295 | AA  | CAATGCTTGAAGATTTTCAACATCT  | TGAGTAAACATGCTTGAACCA    |                          | 3354 |
| D | 5840 | AA  | CAATGCTTGAAGATTTTCAACATCT  | TGAGTAAACATGCTTGAACCA    |                          | 5899 |
| O | 3355 | AT  | GAATTTGAGCAGTACTCTCTCTCT   | CTCTCTCTCTCTCTCTCTCTCTCT |                          | 3414 |

---

|   |      |    |                            |                          |           |      |
|---|------|----|----------------------------|--------------------------|-----------|------|
| D | 5900 | AT | GTGACCTGGAAAAAGATGCTGAC    | TGAGCCTGATTTGACCCCTTGCT  | GCTGCCACA | 5959 |
| O | 3415 | CT | ACACATGACCCCTGCTCATGAGACA  | AGTACAGTACAGAAATTCGCTCT  | TTT       | 3474 |
| D | 5960 | CT | ACACATGACCCCTGCTCATGAGACA  | AGTACAGTACAGAAATTCGCTCT  | TTT       | 6019 |
| O | 3475 | TC | ACATCTTTGATGAGACCAAAAGCT   | GTACTTCACTGAAATATGAAAGAA | CTGCA     | 6079 |
| D | 6020 | TC | ACATCTTTGATGAGACCAAAAGCT   | GTACTTCACTGAAATATGAAAGAA | CTGCA     | 6079 |
| O | 3535 | GG | GTCTCCCTGCAATATGCAATGGA    | ATGATCCACTTTTAAAGAAATAT  | GCTTCATG  | 6139 |
| D | 6080 | GG | GTCTCCCTGCAATATGCAATGGA    | ATGATCCACTTTTAAAGAAATAT  | GCTTCATG  | 6139 |
| O | 3595 | CA | ATCAATGCTTACATATGATGATGAT  | ACTACTGCTTATGATGCTCAGAT  | CAAAAGCA  | 6199 |
| D | 6140 | CA | ATCAATGCTTACATATGATGATGAT  | ACTACTGCTTATGATGCTCAGAT  | CAAAAGCA  | 6199 |
| O | 3655 | TC | TCGATGGAACCCAGGCTATGATTTCT | GGGTCCTCAGACTCTTGGAACA   |           | 3714 |
| D | 6200 | TC | TCGATGGAACCCAGGCTATGATTTCT | GGGTCCTCAGACTCTTGGAACA   |           | 6259 |
| O | 3715 | GA | ATGCTTGAAGATTTTCAACATCT    | TGAGTAAACATGCTTGAACCA    |           | 3774 |
| D | 6260 | GA | ATGCTTGAAGATTTTCAACATCT    | TGAGTAAACATGCTTGAACCA    |           | 6319 |
| O | 3775 | CA | ATGCTTGAAGATTTTCAACATCT    | TGAGTAAACATGCTTGAACCA    |           | 3834 |
| D | 6320 | CA | ATGCTTGAAGATTTTCAACATCT    | TGAGTAAACATGCTTGAACCA    |           | 6379 |
| O | 3835 | GC | CTTATGAGCAGTACTCTCTCTCTCT  | CTCTCTCTCTCTCTCTCTCTCTCT |           | 3894 |
| D | 6380 | GC | CTTATGAGCAGTACTCTCTCTCTCT  | CTCTCTCTCTCTCTCTCTCTCTCT |           | 6439 |
| O | 3895 | AG | TCGATGGAACCCAGGCTATGATTTCT | GGGTCCTCAGACTCTTGGAACA   |           | 3954 |
| D | 6440 | AG | TCGATGGAACCCAGGCTATGATTTCT | GGGTCCTCAGACTCTTGGAACA   |           | 6499 |
| O | 3955 | CA | GAATTTGAGCAGTACTCTCTCTCT   | CTCTCTCTCTCTCTCTCTCTCTCT |           | 4014 |
| D | 6500 | CA | GAATTTGAGCAGTACTCTCTCTCT   | CTCTCTCTCTCTCTCTCTCTCTCT |           | 6559 |
| O | 4015 | AT | GAATTTGAGCAGTACTCTCTCTCT   | CTCTCTCTCTCTCTCTCTCTCTCT |           | 4074 |
| D | 6560 | AT | GAATTTGAGCAGTACTCTCTCTCT   | CTCTCTCTCTCTCTCTCTCTCTCT |           | 6619 |
| O | 4075 | TT | ATTCACGGCATGAGACCCAGGCTCT  | GGATGCAAGGCTGATGCTTGCC   | ACCATCA   | 4134 |
| D | 6620 | TT | ATTCACGGCATGAGACCCAGGCTCT  | GGATGCAAGGCTGATGCTTGCC   | ACCATCA   | 6679 |
| O | 4135 | AG | TTATCATCATGATGATGATGATGAT  | GATGATGATGATGATGATGATGAT |           | 4194 |
| D | 6680 | AG | TTATCATCATGATGATGATGATGAT  | GATGATGATGATGATGATGATGAT |           | 6739 |
| O | 4195 | CT | GAATTTGAGCAGTACTCTCTCTCT   | CTCTCTCTCTCTCTCTCTCTCTCT |           | 4254 |
| D | 6740 | CT | GAATTTGAGCAGTACTCTCTCTCT   | CTCTCTCTCTCTCTCTCTCTCTCT |           | 6799 |
| O | 4255 | TT | ATTCACGGCATGAGACCCAGGCTCT  | GGATGCAAGGCTGATGCTTGCC   | ACCATCA   | 4314 |
| D | 6800 | TT | ATTCACGGCATGAGACCCAGGCTCT  | GGATGCAAGGCTGATGCTTGCC   | ACCATCA   | 6859 |
| O | 4315 | GC | ATCTTGAAGATTTTCAACATCT     | TGAGTAAACATGCTTGAACCA    |           | 4374 |
| D | 6860 | GC | ATCTTGAAGATTTTCAACATCT     | TGAGTAAACATGCTTGAACCA    |           | 6919 |
| O | 4375 | TC | TCGATGGAACCCAGGCTATGATTTCT | GGGTCCTCAGACTCTTGGAACA   |           | 4434 |
| D | 6920 | TC | TCGATGGAACCCAGGCTATGATTTCT | GGGTCCTCAGACTCTTGGAACA   |           | 6979 |
| O | 4435 | TC | TCGATGGAACCCAGGCTATGATTTCT | GGGTCCTCAGACTCTTGGAACA   |           | 4494 |
| D | 6980 | TC | TCGATGGAACCCAGGCTATGATTTCT | GGGTCCTCAGACTCTTGGAACA   |           | 7039 |



|    |      |  |      |
|----|------|--|------|
| OY | 4495 | gacctcaggtgnaaatalccaaaagatgycgtgcaaglygactccagaagaacatlyaaag  | 4554 |
|    |      |  |      |
|    |      |  |      |
|    |      |  |      |
|    |      |  |      |
| Db | 7040 | GACCTCAGGTGAATAATCCAAAAGAGGCGTCAATGGACTTCCAGAAAGCAATGAAG       | 7099 |
|    |      |  |      |
| OY | 4555 | tcaacaggaatfaactactccaaaggagtaaatctctgcttccaaagatfatytgaaagat  | 4614 |
|    |      |  |      |
| Db | 7100 | TCACAGGAGTAACTACACTCAGGGAGTAAATCTCTGCTTACGACAGATGATGTGAAGAGT   | 7159 |
|    |      |  |      |
| OY | 4615 | tcccatctccagagatcaaatgagcatcatagtgagctctcttttgaatgcaag         | 4674 |
|    |      |  |      |
| Db | 7160 | TCTCATCTCCAGCAGCAAGATGGCCATCAGGGACCTCTTTTTCAGATGGCAAG          | 7219 |
|    |      |  |      |
| OY | 4675 | taaaagtcttccaaaggaaatcaagactccctcaacctgtgtgaacctctaaagccac     | 4734 |
|    |      |  |      |
| Db | 7220 | TAAAGGTTTTCAAGGGAATCAAGACTCCTTCACACTGTGTGTAACCTCTAGACCCAC      | 7279 |
|    |      |  |      |
| OY | 4735 | cgttactgaactcgtactactcgaattcacccccaagatgtgggtgacagaaattgcctcga | 4794 |
|    |      |  |      |
| Db | 7280 | CGTTACTGACTCGGTACTACTTGAATTCACCCCAAGTGGGTGACCAAGATGGCCCGA      | 7339 |
|    |      |  |      |
| OY | 4795 | ggatgagaggtctcgagctcgcagagcaagagacctactgaactcgaagcagatc        | 4848 |
|    |      |  |      |
| Db | 7340 | GGATGAGAGTTCGGCGCTGGAGGCAACAGACTCTACTAGAGAGGTGGCCACTGCG        | 7393 |
|    |      |  |      |



|   |      |  |      |
|---|------|--|------|
| D | 1104 | GAAGAAACCGGAGACTATGATGATGATCTACTGATCTGAAATGAGATGCTCAGGTTT    | 1163 |
| O | 1557 | gaigaiaaacaatcctcctccttataccaatctcaatitgagtgcaaaagaatctaa    | 1616 |
| D | 1164 | GATGATGACAACTCTCTCTTATCCAAATTCGCTCAGTTGCCAAGAGATCTCTAA       | 1223 |
| O | 1617 | acttggttacaatcacttgccttgaagaaggactgagactaagctcccttaagtcctc   | 1676 |
| D | 1224 | ACTTGGGTACTATTCATTTGCTGTGAAGAGAGGACTGGACTATGTCCTTACTGCTC     | 1283 |
| O | 1677 | gccccgatatgcaaatgttataaaagtcaatattgaaacaatggcccccgaggttgg    | 1736 |
| D | 1284 | GCCCCGATGACAGAAAGTTATAAAGTCAATTTGAAACAAAGGCCCTCAGGGATTGGT    | 1343 |
| O | 1737 | aggaagtacaaaagttccgatttattggtacacacagatgaacctttaagactgtgaa   | 1796 |
| D | 1344 | AGGAAGTACAAAAAGTCCGATTTATGGCATACAGATGAACCTTTAAGACTGCTGA      | 1403 |
| O | 1797 | gctattcaagatgaatcagaatccttggaaaccttacttatttggtgaagtltgagacaa | 1856 |
| D | 1404 | GCTATTCAGATGATGATGAGATCTTGGACCTTACTTTATGGGAGTTGGAGACAA       | 1463 |
| O | 1857 | ctgttgattataattagaatcaagcaagcaacacataaactacccccaaggaatc      | 1916 |
| D | 1464 | CTGTGATTATTTAAGAAATCAAGCAAGACCATATTAACATCTACCCCTCAGCAATC     | 1523 |
| O | 1917 | actgabtccgctccttggatttcaagagatctacaaaagttgtaaaacatttgaagt    | 1976 |
| D | 1524 | ACGTATCTCCCTCTTGTATTCANAGAGATTACCAAAAGGTGTAACATTTGAAGAT      | 1583 |
| O | 1977 | tttccaattctcgcaggagaatactcaataataatlgacagtgactgtagaagtgg     | 2036 |
| D | 1584 | TTTCCAAATTCGCCAGGAGAAATTTCAATATTAATGAGACAGTACGTGAGAAATGGG    | 1643 |
| O | 2037 | ccaactaatcaagatcctcgtggtccctgaaccgctactactagtttcgttaataagg   | 2096 |
| D | 1644 | CCAACCTAAATCAGATCTCGTGGCTGACCCGCTAATCTAGTTTGCTTAATTTGAG      | 1703 |
| O | 2097 | agagatctagcttcaagactatctggccccctcctcactcctctacaagaatcttagt   | 2156 |
| D | 1704 | AGGATTTACTAGTTAGGACTCATTTGGCCCTCTCCATCTGCTACAAAGAAATCTAGAT   | 1763 |
| O | 2157 | caaaaggaacaacagataatgtcagaacaagaagaatgtcatcctgtttctgtattgat  | 2216 |
| D | 1764 | CAAAAGAGAAACAGATATATGTCAAGACAAAGAGAAATGTCACTCTGTATTTGAT      | 1823 |
| O | 2217 | gaagaaccgaagctgtgtacctacagagaaataaagaagctttctcccaatccagctga  | 2276 |
| D | 1824 | GAGAAACGGAAGCTGTACTCTACAGAGAAATATACAAAGCTTTCTCCCAATCCAGCTGA  | 1883 |
| O | 2277 | gtcagacttgagatcagagatctccaagccctccaacatcagacagcatcaatgtcat   | 2336 |
| D | 1884 | GTCGAGCTTGAGATCCAGAGATTCCAAGCCTCCAACATCATGACACAGATCATGGCTAT  | 1943 |
| O | 2337 | gtttcttgatagttgtcagctgtgtgttttgatagaggtgacatctgtgtacattcta   | 2396 |
| D | 1944 | GTTTTGATAGTTGTGAGTTGTAGTTGTTTGCATGAGTGGCATCTGGTACATTTCTA     | 2003 |
| O | 2397 | agcaatggagacacagactgtacctctctgtctctctctgtgataactctcaaac      | 2456 |
| D | 2004 | AGATTGAGACACAGACTGACTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT    | 2063 |
| O | 2457 | aaaaatgtctatgagac    | 2516 |
| D | 2064 | AAAATGCTATGAGACACACACACACACACCTATTTCCATTTCTAGAGAAATGTCTTCA   | 2123 |
| O | 2517 | tcgaatggaaaaacccaaggttctatgtgattctgggtgtccacaacacacacacacac  | 2576 |
| D | 2124 | TTCATGTAAGAAACCAAGGCTATGATTTCTGGGAGTCCCAACACTGAGACTTTTCGA    | 2183 |
| O | 2577 | ggcatgacgccttactgaaaggttctagttgtgacaaagaacactgtgtattatagag   | 2636 |

---

|   |      |  |      |
|---|------|--|------|
| D | 2184 | GGCATGACCGCCTTACTGAGAGTTTCTAGTTGTGTGACAAAGAACTGATGATTTACAG     | 2243 |
| O | 2637 | gacagttaagaatatttcaagactactgtgtgagtaaaacaatgtccattgaacaa       | 2696 |
| D | 2244 | GACACTTATGAAGATATTTACAGTACTTGTGAGTGAATTAAGAAATGCAATGAAACAA     | 2303 |
| O | 2697 | agctctcccaaatctcaagaacacacttagcactaggtgcaaaagaatthaatgcccaca   | 2756 |
| D | 2304 | AGCTTCTCCGAAATTCAGACACACCTAGCTAGGCGCAAAAGCAATTTAATGCCACCCA     | 2363 |
| O | 2757 | caagtcttgaagcgcatacaacgggaaaaataactcgttactactcttcaagtgacaaag   | 2816 |
| D | 2364 | CCAGTCTTGAACCGCATCAACGGGAAATTAACGTGTACTCTCTCTCTCTCTCTCTCT      | 2423 |
| O | 2817 | gaatttgacatgatgatacaatcactgattgaatgaagaagaagatttgaacattat      | 2876 |
| D | 2424 | GAAATTTGACTATGATGATACCATATCATTTGAATGAAGAAGAAATTTGACATTTAT      | 2483 |
| O | 2877 | gatgaggaatgaanaatcagaagccccgcagcttcaaaaagaacaagacatattat       | 2936 |
| D | 2484 | GATGAGATGAAAAATCAGAGCCCCCGAGCTTTCAAAAAGAAACAGACACTATTTAT       | 2543 |
| O | 2937 | gtctgagtgagagagctctgggattatgtggtatgtgtcccaatglttcaagaac        | 2996 |
| D | 2544 | GCTGCAGTGAAGGCTCTGGGATTATGGATGATGATCTCCCATGTTCTTAAGAAC         | 2603 |
| O | 2997 | agggtcagagtgagtgagctgcccctcagttcaagaagaatgttttccaagaattacgat   | 3056 |
| D | 2604 | AGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTGTGTTCCAGAAATTTACTGAT    | 2663 |
| O | 3057 | ggctccttactcagcccccttataccgttgagaactaaatgacaatttgggaactctggg   | 3116 |
| D | 2664 | GGCTCTTACTACACCCCTTATACCGTGAGAACATTAATGAACATTTGGAGCTCTGGGG     | 2723 |
| O | 3117 | ccatatataagaagcaagtgatgaagataataatcatgttacttccagaatcagcctct    | 3176 |
| D | 2724 | CCATATATTAAGAGCAAGAGTTGAAGATATATCATATGTAATCTTTCAGAAATCAGGCTCT  | 2783 |
| O | 3177 | gtcccttactccttactcactcacttacttcttataatgagaagaatcagaagcagaagca  | 3236 |
| D | 2784 | CGTCCCTATTCCTTATTTATTTAGCCTTATTTCTTAATGAGAAATCAGAGGAGGACA      | 2843 |
| O | 3237 | gaacctagaanaaacttgtcaagcctaaatgaacaaacttacttlttgaaagtgc        | 3296 |
| D | 2844 | GAACCTTGAAGAAATCTTGTCAAGCCTATATGAACCAAAATCTTGTGAAAGTCAA        | 2903 |
| O | 3297 | catcatatggaccccaacaaagatgaatgttgaacgcaaaagccttggttattctctgat   | 3356 |
| D | 2904 | CATCATATGACACCCACTAAAGATGATTTGACTGCAAAAGCTGGCTTATTTCTGAT       | 2963 |
| O | 3357 | gttgacctggaanaaagtgtgcaactcaggttgaacccctctgtgtcgcacact         | 3416 |
| D | 2964 | GTTGACCTTGGAAGAAATGTGCACACTCAGGCTGATTTGAGACCCCTTGTGTCACACT     | 3023 |
| O | 3417 | aaacacatgaacccctgtcatalggtgagacaagtgacagtlacagaatltgctgttttctc | 3476 |
| D | 3024 | AAACACACTGAACCCGTGCTCAGTGGAGACAACTGACAGTACAGGAATTTGCTGTTTTTC   | 3083 |
| O | 3477 | aacatcttltgtatgaccaaagaactgtgacttaactgaataatggaagaagaacttgag   | 3536 |
| D | 3084 | ACCATCTTGTATGAGACCAAAAGCTGTACTCTGCTGAAATATGGAAGAAACCTGACGG     | 3143 |
| O | 3537 | gtcccttgacatatacaagatgaagatcccaactttaagaagaaatlaagcttccatgca   | 3596 |
| D | 3144 | GCTCCCTGCATATTCACATATGGAAGATGCCACTTTTAAGGAATTTATCGTTCCATGCA    | 3203 |
| O | 3597 | atcaatgtctacataatgtatatactactgtgttgatlaatgtgtcaagatcaaaagat    | 3656 |
| D | 3204 | ATCAATGCTACTATATGATGATACACTACTGCTTAGTAATGAGCTCAGATCAAAAGAT     | 3263 |
| O | 3657 | cgatgtatctgtctcagatgtggcagcaatgaaanaacatcatctatctcaatctcaagga  | 3716 |
| D | 3264 | CGATGTATTCGTACAGATGGCAGCAATGAAGAAACATTCATTTCTATTTCACTGTGA      | 3323 |

QY 3717 catgttctactgtacgaaaaaagagatataaatgagcactgtacactctatca 3776  
|||||  
Db 3324 CATGTTTCACTGTRACGAAAAAAGAGAGATATAAATGGCACTGTACAATCTCATCA 3383  
QY 3777 ggtgtttttgagacagtgaaatgttaccatcaaaagcttggaattgagcgggtggaatgc 3836  
|||||  
Db 3384 GGTGTTTTGAGACAGTGAATGTTACATTCAAAGCTGGAAATTTGGCGGTGAATGC 3443  
QY 3837 ctatattgagacatctacatgctggtggaatgacacacttttcgtgttacaagaataag 3896  
|||||  
Db 3444 CTATATGGCGAGCATCTACATGCTGGATGAGACACTTTTTCGTGTACAGCAATTAAG 3503  
QY 3897 tgcgaactccccggggaatgtcttcgagacatllaagaatttcagatlaacagcttca 3956  
|||||  
Db 3504 TGTCAAGACTCCCCGTGGGAATGCTTTGTGACACATTAAGATTTTCAGATTACAGCTTCA 3563  
QY 3957 ggaacaatatgacagatggtggtcccaaaagcttgccagacttcatlccggaatcaat 4016  
|||||  
Db 3564 GGACAAATATGAGAGATGGGCCCCAAAGCTGGCCAGACTTCAATTAATCCGATCAATCAAT 3623  
QY 4017 gcttgggagacacaaagagacccctttcttgatcaaatgtgactgttggcaacaaatgatt 4076  
|||||  
Db 3624 GCTGGAGACCAAGAGACCCCTTTTCTTGATCAAGGTGATCTGTGGCAACAAATGAT 3683  
QY 4077 atlcagagcatcaagacccagatggtcccgatcaagaagtctcagccctctacatctcag 4136  
|||||  
Db 3684 ATTACAGCATCAAGACCCAGAGGTGCCCTGACAGAAATTTCTCAGACTCTACATCTCTAG 3743  
QY 4137 ttatcatcatglatagctctgatgtggaagaagtgcagacttcatcgaagaattccact 4196  
|||||  
Db 3744 TTATATCATCATGATAGTCTTGATGGAGAGAAAGTGGAGACTTATFCAGAAATTTCCACT 3803  
QY 4197 ggaaccttaatggtctctcttttgcaaatgtgattcatcttggaataaacaataatttt 4236  
|||||  
Db 3804 GGAACCTTAAGGTCTTCTTGCAATGTGATTCATCTGGGATAAACAACAATATTTT 3863  
QY 4257 aacctccaattatctgcgtacatccgttttgcaacccaactcatlaagacttcagc 4316  
|||||  
Db 3864 AACCTCCAATTAATGCTCGATACATCCGTTTGCAACCACTCATTAAGATTTCCAGC 3923  
QY 4317 accttcgcagatgagatgagctggtgatttaataatgttcgaacatgccaattggaatg 4376  
|||||  
Db 3924 ACTCTTGCGATGAGATGATGGCTGTGATTAATAATAGTGCAGATGCCATTTGGGAATG 3983  
QY 4377 gagagtaagaataataataatgacagagattacgtcttcatctacttaacaataattt 4436  
|||||  
Db 3984 GAGAGTAAGCAATATACAGATGACAGATTAAGTCTTCACTTACCAATATGTTT 4043  
QY 4437 gcaacctgtctccttcaaaagctcgaacttcaactccaagaaggaatgaatgctgga 4496  
|||||  
Db 4044 GCCACCTGCTCTCTTCAAAAGCTCGACTTCACTCCAAAGGAGAGTAATGCTGGAGA 4103  
QY 4497 cctcaggtgaaataatcaaaagatggtcgtgaagtgtgacttccgaagaacatgaaatg 4556  
|||||  
Db 4104 CCTCAGGTGAATATATCAAAAGAGTGGCTGCAAGTCCAGTCCAGAAACATGAAATC 4163  
QY 4557 acagagatgaactactcaagggggtataaatctctgcttaccagcatgtatgtgaaggatgc 4616  
|||||  
Db 4164 ACAGAGTAATACTCTGAGGGGTAAATCTGCTTACCAACATGATGTCAAGAGATTC 4223  
QY 4617 ctcatctccagcagatcaagatggtcagcatcagtgagactctcttttcagaatggaagaatg 4676  
|||||  
Db 4224 CTCACTTCACAGACTCAAGATGGCCATCAGTGAATCTCTTTTTCAGAAATGGCAAAATGA 4283  
QY 4677 aaggtttttcagggaaatcaagactccttcaacactgtgtgaactctctlaagaccacg 4736  
|||||  
Db 4284 AAGSTTTTTCAGGAAATCAAGACTCTCTTACACTGTTGGTGAATCTCTAGACCCACCG 4343  
QY 4737 ttactgactgctactccttggaattcaacccccagagttgggtgtgacccagattgcccctgagg 4796  
|||||  
Db 4344 TTACTGACTGCTACTCTTCGAATTCACCCCCAGAGTTGGGTGACACAGATTTGCCCTGAGG 4403

QY 4797 atgaggtcttcggtcgtcgaggaagcaagacacttactgactcagcgagttc 4848  
|||||  
Db 4404 ATGAGAGTTCTGGCTGCGAGGACACAGSACCTTACTAGAGGTGCGCACTGC 4455

RESULT 8  
US-08-882-083-1  
: Sequence 1, Application US/08882083  
: Patent No. 5869292  
: GENERAL INFORMATION:  
: APPLICANT: VOORBERG, Johannes J.  
: TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY  
: NUMBER OF SEQUENCES: 17  
: CORRESPONDENCE ADDRESS:  
: ADDRESS: Foley & Lardner  
: STREET: 3000 K Street, N.W., Suite 500  
: CITY: Washington  
: STATE: D.C.  
: COUNTRY: USA  
: ZIP: 20007-5109  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/882,083  
: FILING DATE:  
: CLASSIFICATION: 514  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/558,107  
: FILING DATE: 13-NOV-1995  
: ATTORNEY/AGENT INFORMATION:  
: NAME: ISACSON, John P.  
: REGISTRATION NUMBER: 33,715  
: REFERENCE/DOCKET NUMBER: 30472/212  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (202)672-5300  
: TELEFAX: (202)672-5399  
: TELEX: 904136  
: INFORMATION FOR SEO ID NO: 1:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 5035 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: Linear  
: FEATURE:  
: NAME/KEY: CDS  
: LOCATION: 35..5017  
: US-08-882-083-1

Query Match 47.2%; Score 3748; DB 2; Length 5035;  
Best Local Similarity 87.4%; Pred. No. 0;  
Matches 4374; Conservative 0; Mismatches 60; Indels 570; Gaps 2;

QY 415 ccacacatgcaaatagagctctccacctgctctcttctgtgctcttgcattcgcttta 474  
|||||  
Db 30 CCACCATGGAATAGAGCTCCACCTGCTTCTGTGCTTTGGATTCGCTTTA 89  
QY 475 gtgcacacagaagatactactgtggtgagtggaactgtcatggaacataatgcaaatg 534  
|||||  
Db 90 GTGCCACCAAGAAATACACTGCTGGGTGAGTGAACGTCTCATGGACTATATGCAAAAGTG 149  
QY 535 atctcgtgtaagctgctgttggaacgaagattccctctagagtgccaaatcttttccat 594  
|||||  
Db 150 ATCTGCGTGAAGCTGCTGTGAGCAGCAAGATTTCTCTAGAGTGCAGAAATCTTTTCAT 209  
QY 595 tcaaacactcagctggtgtacaaaaagactcgtgtttagaattcaaggttcaacttttca 654  
|||||  
Db 210 TCAACACCTCAGTGTGTACAAAAAGACTGTGTTGTAAGATTTACCGGATCACCTTTTCA 269  
QY 655 acatgcctaaagcgaagccaccctgtgattggtctgtctagtgcttaccatccaggtcagg 714

Db 270 ACATGCGCTAAGGCAAGGCGACCCCTGGATGGGCTGCTAGTGCCTTACCATCCAGGCTGAGG 329  
OY 715 tttaatgatacagttggtcatttaacactaaagaatgctccatccctgacatcttcacg 774  
Db 330 TTTATGATACAGTGCTACCTTACCTTAAAGACATGGCTTCCATCTGCTGACGTCTTCAAG 389  
OY 775 ctgtctgtctatccctactggaagactctctgagagagcttgatatgatagcacacagtc 834  
Db 390 CTGTTGGGTATCTTACTAGTGAAGAGCTCTGAGGAGCTGAATATGATGATCAGACAGTC 449  
OY 835 aagaaggagaagaatgataaagctctccctggttggaaagccatatactgtctggcag 894  
Db 450 AAAAGGAGAAAGAAAGATATAAGTCTTCCCTGGTGAAGCCATACATATGCTGGCAGG 509  
OY 895 tccgaaagagaatggtccaatggtcctgacccactggtcccttaccctatatactt 954  
Db 510 TCCGTGAAGAGATGCTCAATGAGCTCTGACCCACTGCTTACCTTACTATATCTTT 569  
OY 955 ctcatgtggaacctggttaaaagacttgaaactcagacctcatctggagccctactagta 1014  
Db 570 CTCATGTGAGCTGCTTAAGAGCTGAATTCAGGCTCATTTGAGGCCCTACTAGTATGTA 629  
OY 1015 gagaaggagctctggccaaagaaagacagaccttgacacaaattataactatttg 1074  
Db 630 GAGAGGAGAGTCTGGCCAAAGGAAAGACACAGACCTTGACAAATTTATACATCTTTTG 689  
OY 1075 ctgtatttgatgaaggaaagcttggaactcagaacaaagacctctgtatgacagata 1134  
Db 690 CTGATTTTGAATGAAGGAAAGAAAGTTGGCACTCAGAAACAAAGATCTTGATGACAGATA 749  
OY 1135 ggaagtctgacatctgctcgggacctggcctaaatgacacaaagtcacatggtlta 1194  
Db 750 GGGATGCTGCAATCTGCTGGGCTGGCCCTAAATGACACAGTCAATGATATGTAACA 809  
OY 1195 ggtctctgcaaggtctgatttgatgacacaaagaaatcagtcattggcattgattgga 1254  
Db 810 GGTCTGTCGACGGTCTGATTTGGATGCCACAGGAAATCACTGATTTGGCATGATGGA 869  
OY 1255 tgggaacacacccctgtaagatgacatcaatactcctcgaaggtacacaaattctgtgaga 1314  
Db 870 TGGGACACACCTCTTAATAGCATCAATTTCTCGAAGGTACACATTTCTTGTGAGGA 929  
OY 1315 accatcgccagcgctccttggaatctcgccaaataacttccctactgctcaacaact 1374  
Db 930 ACCATCGCCAGCGCTCTTGGAAATCTGCCAATTAACCTTCTCTACTGCTCAACACCTCT 989  
OY 1375 tgatggaccttggaagtttctactgcttttgcatactcttcccaacaaactgattgga 1434  
Db 990 TGATGAGACCTTGGACAGTTTCTACTGTTTGTTCATATCTCTTCCACCAACATGATGGA 1049  
OY 1435 tgggaactatgctcaaaatagacaagctgcccagagaaaccccaactagaaatga 1494  
Db 1050 TGGAGGCTTATGCTCAAAATAGACAGCTGTCAGAGGAACCCCACTAGCAATGAAAAA 1109  
OY 1495 atgagaagcggaagactatgatagtactactgattctgaaatggaatgtgtcaggt 1554  
Db 1110 ATGAGAGAGCGGAAGACATATGATGATCTTACTGATTTCTGAATTTGATGATGTCAGGT 1169  
OY 1555 ttgatgtgaactcttccctcttaccatccaaatcgctcagtttgccaaagaagatccca 1614  
Db 1170 TTGATGATGACACTCTCTCTCTTATTCCAAAATTCGCTCACTTCCAAAGAGATCTTA 1229  
OY 1615 aaacttgggtacatatactgctgctgaagagagagactggtgactatgctcccttagtc 1674  
Db 1230 AAACCTGGGTACATTTACTTGTCTGTAAGAGAGAGAGACTGGGACTATGCTCCCTTAATGTC 1289  
OY 1675 tgcgccccgatgacagaagtataaaatgcaatattgaaacaaatggccctcagcgatg 1734  
Db 1290 TCGCCCCGATGACAGAGATTATAAAGTCAATATTTGAACAATGGCCCTCAGCGGATGG 1349  
OY 1735 gttaggaagtacaaaaaagctcgattatggtacacacagatgaaccccttaagactgtg 1794  
|||||

---

Db 1350 GTAGGAAGTACAAAAAGTCCGATTTTANGCATFACACAGATGAAACCTTTAAGACTCGTG 1409  
OY 1795 aagcatcaagcatgaaatgaagaaatcttggaaccccttactatgaggaaatggaaca 1854  
Db 1410 AAGCTATTGACATGAATAGGAATCTTGAGACTTTTACTTTATGAGGAAGTGGAGCA 1469  
OY 1855 caactgtgataatlaaagaatcaagcaagcagacatataacatctaccctcaaggaa 1914  
Db 1470 CACTGTTGATTTATTTAAGAAATCAAGCAAGCAGACCATATATACATCTTACCCTCAGGAA 1529  
OY 1915 tcaactgattcgtctcttctgtaatcaaggaatcccaaaaggtgtaaaactttgaag 1974  
Db 1530 TCACGTATGTCGTCCTTGTATTCAGGAGATTAACAAAAGGTGTAACAAATTTGAAGG 1569  
OY 1975 atttcccaatctgccaagagaaatatacaataaataagtgaaatgactgataagaatg 2034  
Db 1590 ATTTTCAATTTCTGCCAGAGAAATTTCAATTAATATGACACTGATGATGAGAAATG 1649  
OY 2035 ggcacaataatcagatccctcggtcgctgacccgcatatactagtttcglttaatalg 2094  
Db 1650 GGCCAACTTAATCAGATCCTCGGTGCTGACCCGCTATTAATCTAGTTCCGTTAATATGG 1709  
OY 2095 aagaagatctagcttcaagactcattggtccctctcctcatctgctacaaagaatctgag 2154  
Db 1710 AGAGAGATCTACCTTCAGGACTCATTTGGCCCTCTCTCATCTGCTTACAAAGAAATCTGAG 1769  
OY 2155 atcaagaagaacacagataatgctcagacaagaagaaatgctcctgcttctctgattg 2214  
Db 1770 ATCAAAAGAGAAACAGATATGTCAGAACAAAGAGAAATGTCATCTGTTTCTGATTTTG 1829  
OY 2215 atggaaccggaagctggttacctcacaagaataatacaagcttctcccaatccagctg 2274  
Db 1830 ATGAGAACCGAAGCTGCTGCTCCACAGAAATATCAACGCTTCTCCCAATCCAGCTG 1889  
OY 2275 gattgagcttgaagatccagagttccaaagcctccaaatattgacagagctaaatgct 2334  
Db 1890 GAGTGAGCTTATGAGATCCAGATTCAGAGCTCCAAACATATGACAGCATATGAGCT 1949  
OY 2335 atgcttctgatatgttgacagttgctgagttgcttgcatagagtggtgacatgacatc 2394  
Db 1950 ATGTTTATGATGTTTGCAGTTTGCAGTTTGTTCATGAGGTGACATACATGATTC 2009  
OY 2395 taagcatggagacagactgacttccctctgctctctctctctctctctctctctcaaac 2454  
Db 2010 TAAGCATTTGAGACACAGACTGACTTCTTCTGCTTCTCTGAGATATCTTCAAAC 2069  
OY 2455 acaaaatggtctatgaagacacactcaacctatccacttccagagaaactcttctca 2514  
Db 2070 ACAAATGCTTATGAGACACACTCACCTATTCCTCATCTCAGAGAAATGCTTTCA 2129  
OY 2515 tctcgtggaanaaccagatctcatgattctggtgtgcacaactcagacttcggaaca 2574  
Db 2130 TGTGATGTAAGAAACCAAGGTATGATTTCTGGGTGACAACTCAACTTTGGAAACA 2189  
OY 2575 gaagcatgacgcttactgaaagttctagtgtgacaaagaacactggtgtattatag 2634  
Db 2190 GAGGATGACCCCTTACTGAGGAGTTTCTAGTTTATTCAGAGGGGAGGAGAGACGCG 2249  
OY 2635 aggaacgtatgaagatattca-----gcatctgtgtgataaaca 2679  
Db 2250 ACTATCTGACCTGGAGAGATTTCAATGGAAGACAGACACTACATCAATCTGTCGACA 2309  
OY 2680 atgcacatgaaacaaagaagctctcccaagatccaagaacaccttagcactagcaaaagc 2739  
Db 2310 GTCGTATTTAACAAGAAAGCTTCTCCAGAAATTAAGACACCTTAGCACTAGGCAAAAGC 2369  
OY 2740 aatttaatgcaac----- 2752  
Db 2370 AATTTAATGCCACCAATTCAGAAATGACATAGAAAGACTGACCTGTTTCAC 2429  
OY 2753 ----- 2752  
Db 2430 ACAGAACACTATGCTTAAATATCAAAATGTCCTCTAGTGAATTTGTTGATGCTTGC 2489

QY 2753 ----- 2752  
Db 2490 GACGAGTCTACTCCACATGSGGCTATCCTTATCTGTCTCCAAAGAACCAATATGAGA 2549  
QY 2753 ----- 2752  
Db 2550 CTTTTCTGATGATCCATCAGCTGAGCAATAGACAGTAATAACAGCGCTGTGTAATGA 2609  
QY 2753 ----- 2752  
Db 2610 CACACTTCAGGCCACAGCTCCATCCACAGTGGGACATGGATTATTTACCCCTGAGTCAGGCC 2669  
QY 2753 ----- 2752  
Db 2670 TCCATTAAATTAATGAGAACTGGGGACAACTGCAGATCCTCTTGGCTTGGGATTACC 2729  
QY 2753 ----- 2752  
Db 2730 ACTATGCTACTCAGATPACCAAAAGAGTGGAAATCCCAAGAGAGTACACGAAAAA 2789  
QY 2753 ----- 2752  
Db 2790 CAGCTTTAAGAAAAAGATACCATTTTGTCCCTGAACGCTGTGAAGCAATCATGCAA 2849  
QY 2753 ----- 2752  
Db 2850 TAGCAGCAATAATGAGGAGCAAAATTAAGCCCAAAATAGAGTCACCTGGGCAAGCAAG 2909  
QY 2753 ----- 2752  
Db 2910 GTAGGACTGAAAGCTGTGCTCTCAAAACCACAGCTTGAAAGCGCATCAACGGGAAA 2969  
QY 2785 taactcgtactactcttcacgtcagtlcaagaagaatctgctatgataacatatcag 2844  
Db 2970 TAACCTGCTACTCTTTCACGTCAATCAAGAGAAATGACTATGATGATACATATACAG 3029  
QY 2845 ttgaatgaagaagaatttgacattatgatgaga tgaanaatcagagcccccga 2904  
Db 3030 TTGAAATGAAGAAAGAAATTTTGACATTTATGATGAGATGAAATCAAGCCCCGCA 3089  
QY 2905 gcttcanaaagaacaacgacactatttaltgctgcagtgagaggtcttgagataltg 2964  
Db 3090 GCTTTCAAAAGAAACACGACACTATTTTATGTGTCGACAGTGAAGGCTCTGGGATTATG 3149  
QY 2965 ggatgagtagctcccaatgctcctaagaacaaggtctcagagtgcagtgctccatg 3024  
Db 3150 GGATGAGTACTCTCCCACTGTTCTAAGAAACAGGGCTCAGAGTGGCAGTCCCTCAGT 3209  
QY 3025 tcaagaagaatgtttcccaagaatttactga tggctcccttactcagccctataccgtg 3084  
Db 3210 TCAAGAAAGTTGTTTCCAGGAAATTACTGATGAGCTCTTACTCAGCCCTTATACCCTG 3269  
QY 3085 gagaactaaatgaacatttgagactcctcgggscatatataagaagcagaagtgaagata 3144  
Db 3270 GAGAACTAAATGAACATTTGGGACTCTCGGGGCATATATAAGGCAAGAGTTGAAGATA 3329  
QY 3145 atatcatgtgtaacttccagaatccaggtcctcgtccctatcccttcatctcagcctta 3204  
Db 3330 ATATCATGTGTAACCTTCAGAAATCAGGCCCTCTGCTCCATTCCTTATCTTAGCCCTTA 3389  
QY 3205 ttctctatgagagaatcagaagcagaagcagaacctaagaanaaacttgtaaacgcta 3264  
Db 3390 TTTCTTATGAGGAGATCAGAGGCAAGGACAGAACCTAGAAAAAATTTGTCAAGCCTTA 3449  
QY 3265 atgaanaacaaactaacttctggaagtgcaacatcatatgycaccacctaagaatgagt 3324  
Db 3450 ATGAATAACCAAACTACTTTTGGAAAGTCAACATCATATATGCAACCACTAAAGATGAGT 3509  
QY 3325 ttgacgtcaaacgttggtctatcttctctgagtgtgacctggaaaaagatgycactcag 3384  
Db 3510 TTACTGTCAAAGCCTGGCTTATTTCTGTGATGTGACCTGGAAAAAGATGTGCTCAG 3569

QY 3385 gctcgtatggaccctctctgctcgcacacataacacatgaaacctgctcatgagagac 3444  
Db 3570 GCCGATTGGACCCCTCTGCTGTGGCACACTAACACAGTAACCCGTGCTATGGGAGAC 3529  
QY 3445 aagtgacagtaacaggaattgctcgtcttctcaacalcctltgtagaaccaaaagctggt 3504  
Db 3630 AAGTGACAGTACAGGAATTTGCTGTTTTCACCATCTTTGATGAGACCAAAAGCTGGT 3689  
QY 3505 acttcaactgaaatatggaagaagaactgcaggtctccctcgaatatccaatggaagatc 3564  
Db 3690 ACTTCACTGAAATATGGAAGAAACATGACAGGGCTCCCTCAATATCCAAATGGAAATC 3749  
QY 3565 ccaactttaaagaagaattatcgcttccatgcaatcgaatgctcacaataatgatacactac 3624  
Db 3750 CCACCTTTAAGAGAAATATCGCTTCATGCAATCAATGGCTACATATATGATACACTAC 3809  
QY 3625 ctggcttagtaatggtcacaagatcaaaagatcgaatgta tctgctcagcatggtgacga 3684  
Db 3810 CTGGCTTAGTAATGGCTCAGAGTCAAGATTCGATGATGTATGCTCAGCATGGCGACGA 3869  
QY 3685 atgaanaacatccatcttcatctatcttcaatgagatggttccactgtcagaaaaaagag 3744  
Db 3870 ATGAAAAATGCATCTTATTCATTTCAAGTGAACATGTGTTCACTGTACGAAAAAAGG 3929  
QY 3745 agtaaaatggaacgtgtaacatctctatccagtggttcttgagacagtggaatgttac 3804  
Db 3930 AGTATAAATGGCACTGTACAAATCTCATCCAGGTGTTTGGACAGTGGAAATGTTAC 3889  
QY 3805 catccaagcttgaatltggcgggtggaatgacctatltggcagacatcatatgcttggga 3864  
Db 3990 CATCCAAGCTGGAATTTGGCGGTGGAATGCCATTATTTGGCGAGCATCTCATGCTGGGA 4049  
QY 3865 tggagcaacttcttctggtgtacaagcaataagtgtaagatcccttgaggatggtcttg 3924  
Db 4050 TGAGCACACTTTTCTGTGTGTACAGCAATTAAGTGTGAGCTCCCTGGGAATGGCTTCTG 4109  
QY 3925 gacaacatgaagaatttcaagattcaagcttcaagacaata tggagccccaaagc 3984  
Db 4110 GACACATTAAGATTTTCAATTAACAGCTTCAGGACAAATTTGACAGTGGGCCCAAGC 4169  
QY 3985 tggcaggaacttcatltccggatccaatgcaatgacctggagcaccaaagagcccttctt 4044  
Db 4170 TGCCAGACTTCAATTATTCGGATCAATCAATGAGCTGTGAGCACCAAGAGACCCCTTTCTT 4229  
QY 4045 ggatcaagtggtgactggttgagccaatgattatcaaggtcaaaagaccaggtgccc 4104  
Db 4230 GGATCAAGGTGGATCTGTTGGACCAATGATTTATTCAGGCAATCAAGAACCCAGGCTGCC 4289  
QY 4105 gtcagaagttctcgaagcctctacatcctcagtttcatcatcatgataagcttgatggga 4164  
Db 4290 GTCAAGAGTTCTCCAGCCTCTTACATCTCTCAGTTTATCATCATGATATAGCTTGATGGGA 4349  
QY 4165 agaagtggcagacttalcgaagaatccacitggaaccttaagtgctctcttggcaatg 4224  
Db 4350 AGAAGTGGCAGACTTATTCGAGGAATTTCCACTGGAACCTTAATGAGTCTTTGGCAATG 4409  
QY 4225 tggatcatcttggtgaataaacaacataattttaacctccaattatgtctgataatcc 4284  
Db 4410 TGGATTCATCTGGGAATPAAAAACAATAATTTTAAACCTCCCAATTAATTTGCTCGATACATCC 4469  
QY 4285 gtttgcaaccaactcatatagcatcgcagacactccttgcaaggttgatggagctgtg 4344  
Db 4470 GTTTGCACCAACTCATTAATAGCATTCGCGACCTCTTCCATGAGATGATGATGGCTGTG 4529  
QY 4345 atttaaatgltgcaagcatgcca tgggaatggaaglaaagaacatcatcagatgacaga 4404  
Db 4530 ATTTAATATAGTTCACAGATCCCATTTGGAATGAGAGTAAAGCAATATCAGATCAGACA 4589  
QY 4405 ttactgtctcatcctacttcaacataatgtttgcaacctggtctccttcaaaagctcgac 4464  
Db 4590 TTACTGCTTCATCTACTTTTACCAATATGTTTGCACCTGGTCTCCCTTCAAAAGCTGCAG 4649  
QY 4465 ttcaactccaagggaggtlaatgacctggaacctcgaagtgaataatccaagaagtggc 4524

|||||  
Db 4650 TTCACTCCAGGAGAGATGATCTGAGACCTCAGGTGAATATCAAAAGAGTGGC 4709  
Qy 4525 tgcgaagtgagctcccaagaacatgaagaatcaagaagtagtaacttaactaaggagtaaat 4584  
Db 4710 TGCAGTGAAGTCTCCAGAGCAATGANAATGACAGAGTAATCTAGGAGAGTAAT 4769  
Qy 4585 ctctgctaccagcatglatgtgaaggaggttcctcatalctcaagcagatcaagaatgagcaltc 4644  
Db 4770 CTGCTGTACAGCATGATGATGGAAGGAGTCCATCATCTCCAGAGTCAAGTGGCCATC 4829  
Qy 4645 agtggagctctcttttccagaatggaagttaaggttttccagggaatcaagaatcctc 4704  
Db 4830 AGTGAAGTCTCTTTTTCAGAAATGCAAAAGTAAGGTTTTCAGGGAATCAAGACTCCT 4889  
Qy 4705 tcaaaccttgatgaactctctagaaccacgcttactgactcgcctactctgaattacac 4764  
Db 4890 TCACACTGTGTAAGTCTCTAGACCCAGCTTACTGCTACCTTGAATTCAC 4949  
Qy 4765 cccagagttgggtgacacagatgcccctgagatgagaggttctggtcgtcgaagcacag 4824  
Db 4950 CCCAGATTGGGTGACACAGATTGCCCTGAGAGTGAAGTGTCTGGCTGCGAGGACACAG 5009  
Qy 4825 acctctactgactcgaagcagttc 4848  
Db 5010 ACCTCTACTGAGGTGGCACTGC 5033  
  
RESULT 9  
US-08-558-107-1  
; Sequence 1, Application US/08558107  
; Patent No. 5910481  
; GENERAL INFORMATION:  
; APPLICANT: VOORBERG, Johannes J.  
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/558,107  
; FILING DATE: 13-NOV-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: ISACSON, John P.  
; REGISTRATION NUMBER: 33,715  
; REFERENCE/DOCKET NUMBER: 30472/212  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5035 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 35..5017  
; US-08-558-107-1

Best Local Similarity 87.4%; Pred. No. 0;  
Matches 4374; Conservative 0; Mismatches 60; Indels 570; Gaps 2;  
Qy 415 ccacatcgaatagagctccacccgtctcttctgtgctcttgcgattcgtctta 474  
Db 30 CCACCATGAAATAGTCTCCACCTGCTCTTCTTGTGCTTTTGCATTCTGCTTTA 89  
Qy 475 gtgcaccagaagatctaccctgggtgcagtggaactgctatgggactatatacaaatg 534  
Db 90 GTGCACCAAGAAATACTACTGGGTGAGTGAACGTGTATGAGGATATATGCAAGTG 149  
Qy 535 atctcgtgagctgctgctggaagcaagattcctcctctagaatgagcgaatccttccat 594  
Db 150 ATCTCGGTGAGACTGCTGTGACCAAAATTTCCCTCTAGAGTGCACAAATCTTTTCAT 209  
Qy 595 tcaaacctcagctgctgacaaaagactcgtttgtagaattcaagggttcacctttca 654  
Db 210 TCACACCTCAGTGTGTACAAAAGACTGTTGTGTAATTCAGGATCACCTTTTCA 269  
Qy 655 acatgcctaaagccaaagccctggaatggtctcgtcctagtcctacatccaggtgag 714  
Db 270 ACATGCTAAAGCCAAAGCCACCTGAGTGGTCTGTAGTGTCTACATCCAGGCTGAGG 329  
Qy 715 ttatgatacagtggtcaatacacttaagaacatggtctccatcctgtcaagttcatg 774  
Db 330 TTTATGATACAGTGTGTACTTAAAGAACATGCTCCATCTGTCAGTCTTCATG 389  
Qy 775 ctgttggtgtatcctctacagaaagctctgaggaagctgaatgatgatacaaccagtc 834  
Db 390 CTGTTGGTGTATCCTACTGGAAGCTTCTGAGGAGCTGAATGATGATGACGACGTC 449  
Qy 835 aaaggagaagaagatbataaagctctccctgtgtgaagccatataatgctcgtcag 894  
Db 450 AAAGGAGAAAGAGATATTAATGCTTCCTGTTGGAAGCATACATATGTCTGGCAGG 509  
Qy 895 tctgtaagaagaatggtccaaatggcctcctgaaccatggtcctactactacatcttc 954  
Db 510 TCTGAAAGAGAAATGCTCAATGCTCTGACCCACGCTGCTTACTACTATCTTT 569  
Qy 955 ctcatgtgacctgttaaaagactgaattcagcccatgtgagcctactatgata 1014  
Db 570 CTCATGTGACCTGTGTAAGAACTTGAATTCAGGCTCATTTGGAGCCTTACTAGTATGTA 629  
Qy 1015 gagaaggagctcgtgccaaggaagaagacagacctgtgcacaaattatactactttg 1074  
Db 630 GAGAAGGAGTGTGGCCCAAGGAAAGACACAGACCTTGCAAAATTAFTACTACTTTTG 689  
Qy 1075 ctgtatttgatgaagggaagaagtgtgcaactcagaacaagaactccttgatgcagata 1134  
Db 690 CTGTATTTGATGAAGGGAAGGAAAGTTGGCACTCAGAAACAAACAACTCCTGTATGACGATA 749  
Qy 1135 gggatgctgcatctgctcgtggcctggcctaaatgacacagatcgtatgtataa 1194  
Db 750 GGGATGCTGCACTGTGCTGGGCTGGCCCTTAATAATGCAACAGTAAAGTTATGTANCA 809  
Qy 1195 ggtcctcgtccagtgctgattgtagtgcacaggaatcaagtcattatggcatgtatgaa 1254  
Db 810 GGTCTCTGCCAGTGTGATTGTGATGCCACAGAAATCAGTGTATTTGGCATGTATTGGA 869  
Qy 1255 tgggacacacctcgtgaatgtaactcaatatctcctgaaggtcacacattcctgttga 1314  
Db 870 TGGGACACCACTCGAAGTGCATCAATATTCCTCGAAGGTCACACATTTCTGTGAGA 929  
Qy 1315 accatgcgaagcgtcccttgaaatctgcgaataacttctactgctccaaacactc 1374  
Db 930 ACCATGCGCAGGCGTCTTGGAATCTGCGCAATACCTTCCTTACGTCTCAAAACACTCT 989  
Qy 1375 tgatgaccttggaacagttctactgatttgcatalctctccacacacatgatgca 1434  
Db 990 TGATGACCTTGGAAGTGTCTACTGTTTGTATATCTCTCTCCACACATGATGAGCA 1049  
Qy 1435 tggaaactatgtcaaaatagacagctgtccagaagaaccccaactacgaatgaanaata 1494  
|||||

|    |      |   |      |
|----|------|---|------|
| Db | 1050 | TGNAAGCTTATGTCAAAAGTAGACAGCTGTCCAGAGGAAACCCCAACTACGATGAAAAATA       | 1109 |
| OY | 1495 | atgaagaagcggaaagactatgatgatcttacttgatcttgaatggatggtgcagtt           | 1554 |
| Db | 1110 | ATGAAGAAGCGGAAGACTATGATGATGATGATCTTACGATTTCTGAATGGATGGCGAGGT        | 1169 |
| OY | 1555 | ttgttgatgaaactctctctctcttataccaattgcgtcagttgcgaagaagcatccta         | 1614 |
| Db | 1170 | TTGTGATGATGACCAACCTCCCTCTCTTTATCCAAATTCGCTCACTTTCGCAAGAAGCATCCTTA   | 1229 |
| OY | 1615 | aaacttggtacattacatctgcctcgtcgaagaagaagcctggagacatgctcccttagtcc      | 1674 |
| Db | 1230 | AAACTTGGGTACATTTACATTTGCTGCTGGAAGAGAGAGACTGGGACATATCTCTCTTAGTCC     | 1289 |
| OY | 1675 | tcgcccccgatgacaagaagtataaaagtcacalatttgacaatgycctcagcgcgtatg        | 1734 |
| Db | 1290 | TCGCGCCCGATGACAGAAAGTTATTAAGATCATATTTGTAACATATGGCCCTCAGCGGATGG      | 1349 |
| OY | 1735 | gtatgaagatcaaaaaagttccgatttatgtgcatacagatgaaacctttaagactcgtg        | 1794 |
| Db | 1350 | GTAAGAAGTCAAAAAAGTCCGATTTTATGGCATACAGATGAACCTTTTAAGACTCGTG          | 1409 |
| OY | 1795 | aagctatcagcctgatcgaagaatcttggagcccttacttatabgggaaattggaaca          | 1854 |
| Db | 1410 | AAGCTATTCCAGCATGAATCTGGATTTGGACTTTACTTTATGGGAACTTGTGAACA            | 1469 |
| OY | 1855 | cactgtgatlatatlatlaaagaatcaagcagaagacacalataacatcacctcaaggaa        | 1914 |
| Db | 1470 | CACGTGTGATTTATTTAAGATCAGCAAGCAAGACATATACATCTTACCCTTCAGGAA           | 1529 |
| OY | 1915 | tcactgatgtcggtccctttgtatctcaaggagaatcaacaagaagttaaacatttgaag        | 1974 |
| Db | 1530 | TCACGTGATGTCCCTCTCTTTGATTTACAGAGATTTCCAAAAGGTGTAAACATTTGAAGG        | 1589 |
| OY | 1975 | atttccaattcgcgcaggagaagaatttcaataataatgtaacgtgtactgttaagaagt        | 2034 |
| Db | 1590 | ATTTTCCAAATTTGCCAGAGGAAGATTTTCAAAATATTAATATGACAAGTCACTGTAAACATG     | 1649 |
| OY | 2035 | ggccaacataatcagatctcctcgtgtgcctcgaaccgcctatcacctcagtttcttaatatg     | 2094 |
| Db | 1650 | GGCCAACTAAATGATGATCCGTGGCTCAGCCGCTATTTACTCTAGTTTGGTTAATATGG         | 1709 |
| OY | 2095 | agagagatctagcttcaaggaactcatgtgcccctctcccatcctgctacaagaatctgtag      | 2154 |
| Db | 1710 | AGAGAGATCTAGCTTACAGACATCATATGGCCCTCTCTCATCTGCTACAAAGAAATCTGTAG      | 1769 |
| OY | 2155 | atcaaaaggaaacccagatataatgtaagaagaagaagtatcactcgttcttcgtatttg        | 2214 |
| Db | 1770 | ATCAAAAGAGAAACCAAGATATATCTCAGACAAAGAAAGATCATCTCTGTTTCTGTATATGG      | 1829 |
| OY | 2215 | atgagaaacgaagctgtgatccctccacagaaatatacaagccttctcccaatccagctg        | 2274 |
| Db | 1830 | ATGAGAACCGGAACCTGTGTACCTCAGCAGAAATATACAAACGCTTCTCCCAATCCACTG        | 1889 |
| OY | 2275 | gagtgagccttgtagatccagaagttcccaagcctccacaatcatgcaagaatcaatgctt       | 2334 |
| Db | 1890 | GAGTGCAGCTTGAAGATCCAGAGTTTCCAAAGCCTCCAAATCATGACAGACGATCAATGGCT      | 1949 |
| OY | 2335 | atgtttttttagtagtttgaagttgtcaagttgttcttgcatgaggtggtgcatactcgtgtaactc | 2394 |
| Db | 1950 | ATGTTTTTGTATCTTTTGACAGTTGTCTCAGTTTGTATTCATGAGAGTGGCATATCGGTATCATTC  | 2009 |
| OY | 2395 | taagcatcttgagacaagacttgactcctctctcgtctctctcttcctcggatataccttgaac    | 2454 |
| Db | 2010 | TAAAGCATTTGAGCACAGACTGACTTTCCTTTTCTGTCTTCTCTGATATACCTTCAAC          | 2069 |
| OY | 2455 | acaaaatggtctatgtaagaacacatccacctaattcccaattctcagagaagaactgtctta     | 2514 |
| Db | 2070 | ACAAAATGTCCTATGAAAGACACATCACCTATTCCTCATTTCTCAGGAGAAAGTCTCTTCA       | 2129 |
| OY | 2515 | ttctcgtatggaaaacccaagttctatgtatctctggtgtccacaactcgaactcttcogaaca    | 2574 |
| Db | 2130 | TTGTCGATGAGAAACCCAGAGTCTATGATTTCTGGGTGTCACAACTCAGACTTTCGGAAACA      | 2189 |

|    |      |  |      |
|----|------|--|------|
| QY | 2575 | gaggagatgacgcgccttactggaaggttctagtgtgcgaagaacatggtgatattacg      | 2654 |
| Db | 2190 | gaggatgatgaccccttactggaagggtttctagtgttcacagaggggagagacgcagc      | 2249 |
| QY | 2635 | aggacagttatgaagatatltca-----gcatctgtctgtgtaaaaaa                 | 2679 |
| Db | 2250 | actatcttgagacctggagagaaanatttcagtgaaagacgacctactacatctcgtcgaca   | 2309 |
| QY | 2680 | atgcacatgaaaccaagaagctcttccagaattcgaagacccctagcaatgcgcaaaagc     | 2739 |
| Db | 2310 | gctctatttgaaaccaaagagcttctccagaatttcagaagacaccttagcctagcgaataagc | 2369 |
| QY | 2740 | aatttaagtcgcac-----aatttaagtcgcac-----                           | 2752 |
| Db | 2370 | aattttaatgcccacacacatttcgaaaaattgcacatagacaacactgacccttggttgcac  | 2429 |
| QY | 2753 | -----  | 2752 |
| Db | 2430 | acagaaacacctatgacctaaataatgcataatgtcctctagtattgttgatgctcttgc     | 2489 |
| QY | 2753 | -----  | 2752 |
| Db | 2490 | gacagagtccttactctcacatggcgctatccttattctgattctccaaagaccgaatatgaga | 2549 |
| QY | 2753 | -----  | 2752 |
| Db | 2550 | ctttttctgattgatnccattacctggagcaatagacagtaatatacagcctgtctgaaatga  | 2609 |
| QY | 2753 | -----  | 2752 |
| Db | 2610 | cacacttgagcgccacagctccatcacactgggacatggtatttatccctgagtcagcc      | 2669 |
| QY | 2753 | -----  | 2752 |
| Db | 2670 | ttccaatttaagatttaattagaaaatcggggacaactgcagatcctcttgggtgaatacc    | 2739 |
| QY | 2753 | -----  | 2752 |
| Db | 2730 | actatgtagtcatgattaccaaaagaaagtgsgaaattcccaagagaaagtcacagaaaaa    | 2789 |
| QY | 2753 | -----  | 2752 |
| Db | 2790 | cagcttttaagaaaaagatnaccattttgtccctgaaacgcttgtaaaacaaatcatgcga    | 2849 |
| QY | 2753 | -----  | 2752 |
| Db | 2850 | tagcagcaatataatgaggagacaaaataagcccgaaatagaagtcaactggggcaaaagcag  | 2909 |
| QY | 2753 | -----ccacacagttctgaaagccatacaacgagaaa                            | 2784 |
| Db | 2910 | gttagagctgaaaggctgtgctctcaaaaaccacacagcttgaaaacgcatcaacggcgaaa   | 2969 |
| QY | 2785 | taacctcttactactcttcaagtcagaatcaagaaggaaattgactatgatatacatatcag   | 2844 |
| Db | 2970 | taactcgtactactactctttagtcacagtcacagagaaattgactatgatatacatatcag   | 3029 |
| QY | 2845 | ctgaaatgaaagaagaagattltgacatttga tgaaggatgaaatacaagcccccgca      | 2904 |
| Db | 3030 | ttgaatatgaaagaaagaaatttttgacattttgatgaggaatgaaatatcagagcccccgca  | 3089 |
| QY | 2905 | gcttctcaaaagaaaacaacgacataatttlatctgtcagatggagagctcttggaattatg   | 2964 |
| Db | 3090 | gcttttcaaaagaaaacaacgacataatttttatgtctgcagtggaagcctctggagattatg  | 3149 |
| QY | 2965 | ggatgagtgcctcccaacatgttctcaagaacaagagctcaagatgtagatgtccctcag     | 3024 |
| Db | 3150 | ggatgagtgcctcccaacatgttcttaagaaacacagagctcagagtgtagatgtccctcag   | 3209 |
| QY | 3025 | tcaagaagaatgttcttccagaatltactga tggctcccttactcaagccctataccgtg    | 3084 |
| Db | 3210 | tcaagaagaatgttcttccagaatltactga tggctcccttactcaagccctataccgtg    | 3269 |



QY 3685 gagaactaaatgaacatttggagactcttggggccatataaagagcaagatttgaagata 3144  
| | | | |  
Db 3270 GAGACATGAAGCAATTTGGGACCTCGGGCCCATATATAGAGAGCAAGTTGAAGATA 3329  
| | | | |  
QY 3145 atatcatgttaacttgaagaatcagacctctgcctccattcccttctatcttagactta 3204  
| | | | |  
Db 3330 ATATCATGTGAAGCAATTTGGGACCTCGGGCCCATATATAGAGAGCAAGTTGAAGATA 3389  
| | | | |  
QY 3305 ttcttataagaaagatcagaagcagaagcagaaccttgaaaaaaacttgcgaagccta 3264  
| | | | |  
Db 3390 TTCTTTATGAGGAAGATCAGAGGCAAGGCAAGCAACTGAAATAAACTTTGTCAGAGCCTA 3449  
| | | | |  
QY 3365 atgaacaacaaacttacttcttggaaagtgcacacatcatatggcaccacataaagabagt 3324  
| | | | |  
Db 3450 ATGAACCAACCAATTTACTTTGGAAAGTGCAACATCATATGCGACCCATTAAGATGAGT 3509  
| | | | |  
QY 3325 ttgaactgaaagcctgggactatctctcgaatgttgaaccttgaaaaaaagtgtgcacctag 3384  
| | | | |  
Db 3310 TTGACTGCAAAAGCCTGGGCTTATTTCTCTGATGTGACCTGGAAAAAGATGTGCACTCAG 3569  
| | | | |  
QY 3385 gccctgatcggacccctcttctgtctgcacacacacacacacttgaacctgtctcatctggagagc 3444  
| | | | |  
Db 3570 GCGATGATGGAGCCCTTCTGGCTGCGCAACATACACACTGAACCTGCTCATGGGAGAGC 3629  
| | | | |  
QY 3445 aagtgcacagtcacaggaatttgcctctgttcttccacatcttgaatgagacaaagctgagt 3504  
| | | | |  
Db 3630 AAGTGCAGTACAGAAATTTGCTGTGTTTTCACCATCTTTATGATGAGCAAAAGCTGCT 3689  
| | | | |  
QY 3505 acctcaccgaaataatgaaagaac 3564  
| | | | |  
Db 3690 ACTTCACCTGAAAAATTTGAAAAAGAACTGCAAGGCTCCTCGCAATATCCGATGGAAGATC 3749  
| | | | |  
QY 3565 ccacttttaagaagaatatactgccttccatgcacacacacacacacacacacacacacacac 3624  
| | | | |  
Db 3750 CCACCTTTTAAAGAGAAATATGCGCTTCATGCATCAATCAATGCGTACATCAATGATGATACCTAC 3809  
| | | | |  
QY 3625 ctgctctgaataatggtctcagaatcagaagatcgaatgtatctgtctcagcattgagcagca 3684  
| | | | |  
Db 3810 CTGGCTTAAATGATGAGTCTGAGTCAAGGATTCGATGATGATGATGATGATGATGATGATGATGAT 3869  
| | | | |  
QY 3685 atgaacaac 3744  
| | | | |  
Db 3870 ATGAACACATCATTTCTTATTTCAATTTTCAGTGCATGTGTCTCAGTGCATGCAAGAAAAAGAG 3929  
| | | | |  
QY 3745 agtataaaatgac 3804  
| | | | |  
Db 3930 AGTATAAATGACACTGTACATCTCTATCCAGGTGTTTGTAGACAGTGAATATGTTAC 3989  
| | | | |  
QY 3805 catccaagctcgaagatttggcgggtggaatgccttatgtggcgaacatcacaatgtctggga 3864  
| | | | |  
Db 3990 CATCCAAAGCTGGAATTTGGGGGTGGAATGCGCTTATTTGGGAGCATCTACATGCTGGGA 4049  
| | | | |  
QY 3865 tgaac 3924  
| | | | |  
Db 4050 TGACACACATTTTCTGTGTGTATGACCAATTAAGTGTACAGATCTCCCTGGGAATGGCTTCTG 4109  
| | | | |  
QY 3925 gacacatagaagatttccagattacagcttcaagaacacacacacacacacacacacacacac 3984  
| | | | |  
Db 4110 GACACATTAGAGATTTTTCAGATTACAGCTTCAGACAAATATGAGACAGTGGGCCCAAAAGC 4169  
| | | | |  
QY 3985 tggcgaagactcatatctccggatcaatcaatgtccttgaagcacaagagaccccttcttct 4044  
| | | | |  
Db 4170 TGGCCAGACTTATTTATTCGGAATCAATCAATGCTTGAGGCCCAAGAGGCCCTTTTCTT 4229  
| | | | |  
QY 4045 ggaatcaagtgagatctgttggcacaacatgatattcagaagcacaagagacccaggtgccc 4104  
| | | | |  
Db 4230 GGATCAAGGTGATCTGTGGGACCAATGATTTATTCAGGCAATCAAGACCAAGGCTGCC 4289  
| | | | |  
QY 4105 gtccagaagttctcagaccttcaacatctcagttatcatcatatgatttcttgaatggga 4164  
| | | | |  
Db 4290 GTGAGAGTTCTCCAGGCTTACATCTCTCAGATTATCATCATGATGATGTTGATGGGA 4349  
| | | | |  
QY 4165 agaagtggcagacttatcagaggaatctccactggaaccttaatggtcttcttggcaatg 4224  
| | | | |

Db 4350 AGAAGTGGCAGACTTATGAGGAATTCACATGCAACCTTATGCTCTTCTTGGCAATG 4409  
| | | | |  
QY 4225 tggatcactcgggataaacaac 4284  
| | | | |  
Db 4410 TGGATTCATCTGGATTAACACAAATATTTTAACTTCACATTTTCTCTGATACATCC 4469  
| | | | |  
QY 4285 gtttgcacccaactcatatagacatctgcgaacactcttgcgaatgagtgaatggcctgtg 4344  
| | | | |  
Db 4470 GTTTCACCAACCATCATTTATAGCATTTCCAGACACTTCCGATGGAGTTGATGGGCTGTG 4529  
| | | | |  
QY 4345 atttaaatagttgcagcatgcattgggaatggagagtaaaagcaatatcagatgcacaga 4404  
| | | | |  
Db 4530 ATTTAAATAGTTGACGATGCGATTGGGAATGGAAGTAAAGCAATATCATATCAGACGA 4589  
| | | | |  
QY 4405 ttacagcttcatccttacttaccatlatgttgcgaacctgttcccttcaaaagctgac 4464  
| | | | |  
Db 4590 TTACTGCTTCATCTACTTATACCAATATGTTTCCACCTGGTCTCTTCAAAACCTGCAC 4649  
| | | | |  
QY 4465 ttacacctcaagaggaagtaatgtccttggagacacacacacacacacacacacacacacac 4524  
| | | | |  
Db 4650 TTACCTCCCAAGGAGAGTATGCTGAGACCTCAGGTGAATATCAAAAGAGTGGC 4709  
| | | | |  
QY 4525 tgcagtgagcttccagaagaacatgaatgcacagaggttaacttaacttaaggaataaat 4584  
| | | | |  
Db 4710 TCCAAAGTGCATTCAGAAAGAACATGAAAGTACAGGAGTACTACTAGGGAGTAAAT 4769  
| | | | |  
QY 4585 cctctgtacacagcatgtatgtgaagagtlcctcatctcagcagatcaaatgtgcatc 4644  
| | | | |  
Db 4770 CTCCTGTTACCAAGCATGATGTAAGAGAGTCTCTATCTCCAGCAGTCAATATGCGCATC 4829  
| | | | |  
QY 4645 agtgcactctcttcttcaagaatggcaagaagtttctcagaggaataaagaactcct 4704  
| | | | |  
Db 4830 AGTGCACCTCTCTTTTTCAGATGCAAAAGTAAAGTTTTCAGGGAATCAACATCTCT 4889  
| | | | |  
QY 4705 tcaacacctgtgtgaacctctcgaacccaacgttacttaactcogtacttgaatcacc 4764  
| | | | |  
Db 4890 TCACACCTGTGTGTAACCTCTAGACCCACGCTTACTGCTACTCTTGAATTCAC 4949  
| | | | |  
QY 4765 cccaagtttgggtgcacagatgtccctggagagatggatgttcttggcttgcagagcacag 4824  
| | | | |  
Db 4950 CCCAGAGTTGGGTGCACAGATGCGCTGAGAGTGAAGAGTCTGGGCTGCGAGGACAGG 5009  
| | | | |  
QY 4825 acctctactgactcagcagatc 4848  
| | | | |  
Db 5010 ACCTCTACTGAGGAGGCGCACCTGC 5033  
| | | | |  
RESULT 10  
US-09-243-539-1  
; Sequence 1, Application US/09243539  
; Patent No. 6130203  
; GENERAL INFORMATION:  
; APPLICANT: VOORBERG, Johannes J.  
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/243,539  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:



APPLICATION NUMBER: US 08/558,107  
FILING DATE: 13-NOV-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: ISACSON, John P.  
REGISTRATION NUMBER: 33,715  
REFERENCE/DOCKET NUMBER: 30472/212  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5035 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 35..5017  
US-09-243-539-1

Query Match 47.2%; Score 3748; DB 3; Length 5035;  
Best Local Similarity 87.4%; Pred. No. 0;  
Matches 4374; Conservative 0; Mismatches 60; Indels 570; Gaps 2;

415 ccacatgcaatagagctccacactgctcttctgtagcttgcgattcgttta 474  
|||||  
30 CCACCATGGAATAGAGCTCCACCTGCTTCTGTGCTTTGCGATTCTGCTTTA 89  
475 gtgcaccagaagaatactactggtgagctgagctgcaatggaactatgcaagtg 534  
|||||  
90 GTGCACACAGAAGATACTACTGCTGAGTGAAGTCACTGAGCATATATGAAAGTG 149  
535 atcccgatgagctgctggtgagcgaagattcctcctagagtgccaaatctttccat 594  
|||||  
150 ATCCGATGAGCTGCTGCTGAGCAGCAAGATTCTCTAGAGTGCCAAATCTTTTCAT 209  
595 tcaacacctgctggtgtacaaaagactggtgtgagaatcaaggttcaacctttca 654  
|||||  
210 TCAACACCTGAGTGTGTACAAAAGACTGTTGTAGATTCACGAGATCCTTTTCA 269  
655 acatcgctaaagcgaagccacctgtagtggtctgtaggttccatccatccaggctgag 714  
|||||  
270 ACATCGCTAAGCCAAAGCCACCTGATGGTGTGTAGTCTTACATCCAGGCTGAGG 329  
715 ttatgatacagtggttacttaacttaagaacatggtctccatcctgtaagttcatg 774  
|||||  
330 TTTATGATACAGTGTGCTATTACATTAAAGAACATGAGCTTCCATCTGTGATGCTCATG 389  
775 ctgtgtgtatctacttgtaaaagctctgaaggagctgaatatgatacagaacagtc 834  
|||||  
390 CTGTGTGTATCTCTACTGGAACCTTCTGAGGAGCTGATATGATGATGAGCAGTC 449  
835 aaaggaagaaagaaatgataaagcttccctggtggaagccatatactgctggaag 894  
|||||  
450 AAAGGAGAAAGAAATGATTAAGTCTTCCCTGCTGGAAGCATACATATGCTGCGAGG 509  
895 tccctgaagaagaatggtcgaatggtcctgagccactggtgctactactatatacttt 954  
|||||  
510 TCCCTGAAGAAGAAATGCTCAATGCTCTGACCCAGTGTGCTTCACTACTCATATGTTT 569  
955 ctcatgtgagcctgttaaaagacttgatccaagcctcatctgagagccctactagtgta 1014  
|||||  
570 CTCTATGTGAGACCTGTGTAAGACTTGAATTCAGGCTCATTTGAGCCCTTACTAGTATGTA 629  
1015 gagaaggaagtcctggaaggaagaaagacacagacctgacaaattatactactttg 1074  
|||||  
630 GAGAAGGAGAGTCTGCGCAAGAAAGACACAGACCTTGCACAAAATTATACTACTTTTG 689  
1075 ctgtatttgatgaaggaagaaagttggaactcagaagaagaagaactccttgatgcagata 1134  
|||||  
690 CTGTATTGATGAAGGGAAGTTGGCAGTCAGAAACAAAGAACTCTTGTATGAGAGATA 749

1135 ggaatgctgcaatctgctcgggacctgagcctaaatgacacagatcatggtatgtaca 1194  
|||||  
750 GGAATGCTGCAATCTGCTCGGCTGGCTTAATAATGACACAGTGAATGTATGTAACA 809  
1195 ggtctcgcagagtcgtatgtagtgcacaggaataagtcattatgtgcatgtgtgaa 1254  
|||||  
810 GGTCTCTGCCAGGTCTGTATGATGAGCCACAGGAATCAGTCTATTTGGCATGTGATTGGA 869  
1255 tgggacccaactcctggaagtgcacatcaatatctcgaaggtcacacatcttctgtgaga 1314  
|||||  
870 TGGGACACACCTCGAAGTGCATCAATATCTCTCGAAGGTCACACATTTCTGTGAGA 929  
1315 accatgccaaaggttcttggaaatctcgccaataacttcttactgctcgaacactct 1374  
|||||  
930 ACCATGCCAGGCTCTCTGGAAATCTCGCAATTAATCTTCTTACTGCTTAACACTCT 989  
1375 tgatgaccttgagacagtttctactgtttgtcatatctctccacaacatgatgca 1434  
|||||  
990 TGATGGAACCTTGAGACAGTTCTACTGTTTGTCAATCTCTCCACACATGATGGA 1049  
1435 tgaagccttatgtcaagtagacagctgtccagaagaaacccaactcgaatgaaata 1494  
|||||  
1050 TGAAGCTTATGTCAAGTAGACAGCTGTCCAGAGGAACCCCACTGCAATGAAAAATA 1109  
1495 atggaagaagcggaagactatgatgatgatcttactgattcgaatgatgtgtcaggt 1554  
|||||  
1110 ATGAGAAGCGGAAGACTATGATGATGATCTTACTGATTCGAATGATGATGTGTGAGT 1169  
1555 ttgatgtacaaactctccttcttcttccaaatctgctcagttgccaagaagatccta 1614  
|||||  
1170 TTGATGATGACAACTCTCTCTTATTCAAATTCCTCTGATGCAAGAGATGCTCTA 1229  
1615 aaactggtgtacatatactatgctgtcgtgaagaagagactggaactatgctccctagtc 1674  
|||||  
1230 AAACCTGGGATCAATTAATCTGCTGTAAGAGGAGCTGGGATATGCTCTTATGTC 1289  
1675 tgcgccccgatgacagaagttataaagtcataitltgaaatagggcctcgaagggattg 1744  
|||||  
1290 TCGCCCCGATGACAGAAAGTTATTAAGTCAATTTTGAAACATAGGCCCTCAGCGGATTG 1349  
1735 gtaagaagtaaaaaaagctcgaatttaltggcatatacagaatgaaacttlaagaactcgtg 1794  
|||||  
1350 GTAGAGATGACAAAAGATCGGATTTATGCGATACAGATGAAACCTTTAAGACTGCTG 1409  
1795 aagctatcagcaatgaatcaggaatcttggagccttacttattggygaagttggagaca 1854  
|||||  
1410 AAGCTATTCAAGCAATCAAGAAATCTTGGGACCTTACTTATGAGGGAAGTTGGAGACA 1469  
1855 caactgttatataattaaagaataaagaagaagcaagataaatacttacccttcagaa 1914  
|||||  
1470 CACTGTTGATTTATTTTAAGAAATCAAGCAAGCAACCATTAATATACCTTACCGAA 1529  
1915 tcaatgatlgtcgttcttltlatccaagagatlaaccaaaaggttgaatacatltgaaag 1974  
|||||  
1530 TCACTGATGTCCGTCCTTGTATCAAGAGATTAACCAAAAAGGATTAACATTGGAAG 1889  
1975 atttccaattctgcccgaaggaatatccaatataaatgagcaagtgactgtgaaagt 2034  
|||||  
1590 ATTTTCCAAATTCGCGAGGAATAATTAATAATGAGACAGTACGTGGAAGATG 1649  
2035 ggcacaataatcagaatccctggtgctgagccgctatcttactgaagtttgtaataatg 2094  
|||||  
1650 GGCCAATTAATCAAGATCCCTGCGTGGCTGACCGCTTATCTAGTTTGTTATATG 1709  
2095 agaagatctagcttcaagcatatgacctctcctcactatctgctacaaagaaatctgtag 2154  
|||||  
1710 AGAGAGATGTAGCTTCAAGCATGATGCGCTCTCTCATCTGCTGCAAGAAATCTGTAG 1769  
2155 atcaagaaggaacagataatgtcagaagaaggaatgtcattccctgttttctgtattg 2214  
|||||  
1770 ATCAAAAGAGAAACAGATATGTCAAGACAAAGAGATGTCAATCTGTTTCTGTATTTG 1829

QY 2215 atgagaacgaagctggtacctcacagagaatacaaacgttctcccaatcaagctg 2274  
|||||  
Db 1830 ATGGAACCGAAGCTGTACTCACAGAAATATCAAGCGTTTCTCCCAATCCAGCTG 1889  
2275 gaggcaagctgagagatccagagatcccaagcctccaacatcatgacagatcaatgct 2334  
|||||  
Db 1890 GAGTGCACCTTGAGGATCCAGAGTTCCAAAGCTCCAACTCATGCAAGCATCAATGGCT 1949  
2335 atgtctttagatgttgcagttgctgagtttgcagatgaggtggcagatctgcatatc 2394  
|||||  
Db 1950 ATGTTTATGATAGTTGTCAGTTGTCAGTTGTTGTCATGAGGTGCGCATCTGTCATTC 2009  
2395 taagcattggagacagacgtcactccttctctgtcttctctctcgtgataactccaac 2454  
|||||  
Db 2010 TAAACATTTGGAGCAACAGCTACTCTTCTGCTCTTCTGATATATACCTTCAAC 2069  
2455 acaaaatggtatagaagaacacactcacctcattccattcccaagaagaactgtctca 2514  
|||||  
Db 2070 ACNAATGCTTATGAAACACACTCACCTTATCCATCTCAGAGAACTGTCTTCA 2129  
2515 tgcgatggaacacccaggtctatgaltctgagggtgcacaacactcagacttcggaaca 2574  
|||||  
Db 2130 TGTGATGGAAGAACCCAGCTGTATGATTTGGGGTGCACAACTCAGACTTTCGGAACA 2189  
2575 gaggcatgacgcgtctactgaaagttctagtgtgacaagaacactgtgattatcag 2634  
|||||  
Db 2190 GAGGCAATGACCGCTTACTGAAAGTTTCTAGTTTATTCAGAGGGGAGGAGAGAGAG 2249  
2635 aggcagctatagaagatctca-----gcactactgtctgagtaaaaaa 2679  
| | | | |  
Db 2250 ACTATCTGGACCTGGAGAGATATTCACTGAAAGACGACGACTACATCCAGATCGTCGACA 2309  
2680 atgcattggaacccaagaagctctccccaagatctcaagaacacactcagactaggaacaa 2739  
| | | | |  
Db 2310 GTCTGATTGAACCAAGAGCTTCTCCAGAAATTCAAGACACCCCTAGCTAGCAAAAGC 2369  
2740 aatttaattgcac----- 2752  
| | | | |  
Db 2370 AATTTAATGCCACCAATTCAGAAATGACATAGAGAAGACTGACCTGTTTGAC 2429  
2753 ----- 2752  
Db 2430 ACAGAACACCTATGCTTAAATACAAATGTCTCTAGTATTTGTGATGCTCTTGC 2489  
2753 ----- 2752  
Db 2490 GACAGAGTCTACTTCACATGGGCTATCTTATCTGATCTCCAGAGCAAAATATGAGA 2549  
2753 ----- 2752  
Db 2550 CTTTCTTGATGATCATCACTCGGAGCAATAGACAGTAATAACAGCCTGTCTGAATGA 2609  
2753 ----- 2752  
Db 2610 CACACTTCAGGCCACAGCTCCATCAGAGTGGGACATGTTATTACCCCTGAGTAGGCC 2669  
2753 ----- 2752  
Db 2670 TCCATTTAAGATTAAATGAGAACTGGGAGCAACTGCAGATCCTTCTGTGGATAACC 2729  
2753 ----- 2752  
Db 2730 ACTATGTACTCAGATACCAAAAGAGAGTAGAATCCCAAGAGAGAGTCCACGAGAAAAA 2789  
2753 ----- 2752  
Db 2790 CAGCTTTAAGAAAAAGATACCATTTTGTCCCTGACGCTTGTGAAGCAATCATGTCAA 2849  
2753 ----- 2752  
Db 2850 TAGCAGCAATTAATGAGGACAAATAAGCCGAATAGAGTCACTGTGGCAAGCAAG 2909  
2753 -----ccacaagctctgaaacgcacatcaacgggaaa 2784

Db 2910 GTAGACTGAAAGGCTGTGCTCTCAAAACCCACCAAGTCTTGAAGGCCATCAACGGGAAA 2969  
2985 taactgtaactactctcaagtcagatcaagaaggaattgactatgatatcacaacag 2844  
|||||  
Db 2970 TTAAGTCTGATCTCTTCAAGTCAAGAGGAAATGATGATGATATCATATCAG 3029  
2845 ttgaaatgaagaagaatttgacattatgaltgaaatgaaatcaagaaccccgca 2904  
3030 TTGAATGAGAGAGAAATTTGACATTTATGATGAGAGATGAATAATCAGAGCCCGCA 3089  
2905 gcttccaaagaacacagacactatltatgtcgcagtgagaggtctcggattatg 2964  
|||||  
Db 3090 GCTTTCAAAAGAAACAGACATATTTATGCTGACGTGGAGAGCGCTGGGATATG 3149  
2965 ggaatgagtagctcccccacatgtctcaagaacagggcagaagtggcagtgctcccaag 3024  
|||||  
Db 3150 GGATAGTAGTCTCCCAATGTTCTAAGAAACAGGGCTCAGAGTGGCATCTCCCTCAGT 3209  
3025 tcaagaagttgtttccaggaattactgaatggtcctccttactcaagccctlaaccgtg 3084  
Db 3210 TCAAGAAAGTTGTTCAGGAATTTACTGATGGCTCCTTACTCAGCCCTTATACCGTG 3269  
3085 gagaactaaatgaacatttgggactcctggggccatatataagaagaagttgaagata 3144  
Db 3270 GAGAACTAAATGAAACATTTGGAGCTCGGGGCCATATATAGAGCAGAAAGTTGAAGATA 3329  
3145 atatcatgtiaacttcaagaatcaagcgtcgtccctactcctcctatctactgactta 3204  
Db 3330 ATATATGTTAACTTTCAGAAATCAGGCCCTCTGCTCCCTATCTTATTTCTAGCCCTTA 3389  
3205 ttcttatgagaagatcaagaagcagaagacactagaaaaacttgcacagctca 3264  
Db 3390 TTTCTTATGAGGAAGATCAGAGGACGAGCAGAACCTAGAAAAACTTGTCAAGCCTA 3449  
3265 atgaacacaaaactacttltggaagatgcacatcatatgacccactaaagaatgag 3324  
|||||  
Db 3450 ATGAAACCAAACTTATGTTGGAAAGTCAACATCATATGAGCCCACTTAAATATAGT 3509  
3325 ttgactgcaaacctgggctatctctctgattgttgcactgaaagaaatgtgcacacag 3384  
Db 3510 TTGACTGCAAAAGCCGTGGCTTATTTCTGATGTTGAGACTGGAAAAAGATGTGCACCTAG 3569  
3385 gcttgaattggacccctctgtctgtccacactaaacactgaaccctgtcatatggagac 3444  
Db 3570 GCCTGATTGGACCCCTTGTGCTGCCACACTAACACACTGAACCCGTCTCATGGAGAC 3629  
3445 aagtgacagtcagaatltgtctgttcttccacatcttggatgaaaccaaagctggt 3504  
Db 3630 AAGTGACAGTACAGGAATTTGCTGTTTTCACCATCTTGTGATGAGACCAAAAGCTGCT 3689  
3505 acttcaactgaaataatgaaagaacactgcagggctccctgcaatatcagaatgaatc 3564  
Db 3690 ACTTCACTGAAATATGAGAAAGAACTCAGAGGCTCCCTGCAATATCCAGATGGAATC 3749  
3565 ccaactttaaagagaattatcgtctccatgcaatcaatgylacataatgataactac 3624  
Db 3750 CCACCTTTAAAGAGAAATATGCTTCATGCAATCAATGAGTATGATGATGATGATGATGATG 3809  
3625 ctggcttagtaatggtcagaatcaagaagatctgaatgtatctgtcgaagatgagagca 3684  
Db 3810 CTGGCTTAGTATGCTAGATCAAAAGATTCAGATGATCTGCTCAGATGGGAGCA 3869  
3685 atgaacacatcatctcatcttcaagtgagactgttcaactgtatcgaaaaaaagag 3744  
Db 3870 ATGAAACATCATCTTCTATTTCAATTCAGTGCACATGTGCTTACGTGATCGAAAAAGAGG 3929  
3745 agtataaaaatggcactgtacaaatcctatccaaagtgcttttgaagacagtgaaatgttac 3804  
Db 3930 AGTATTAATGAGCACTGTACATCTTATCCAGGTGTTTGAAGACAGTGAATGTATC 3989  
3805 catcacaagctggaattggcggtggaatgacctatgtgcagacatcatatgcttggga 3864  
|||||

Db 3990 CATCAAGCTGCAATTTGGCGGTGGAAATGCTTATTTGGCGAGCATCTACATGTGGGA 4049  
QY 3865 tggagcaactttcttggtgtacagaataaagtgtaagctccctgggaatgcttcg 3924  
Db 4050 TGAAGCACTTTTCTGGTGTACAGCAATAGTGTAGACTCCCTGGGAAATGGCTTCTG 4109  
QY 3925 gacacattagaatttcagatattacagcttcagagacaatatggaagtgggcccaagc 3984  
Db 4110 GAACATTTAGAGATTTTTCAGATTACAGCTTCAGAGACAATTTGACAGTGGGCCCAAGC 4169  
QY 3985 tggcgaacttcattatcccgatcaatcaatgaagctgagagaccgaagagcccttcct 4044  
Db 4170 TGGCCAGACTTCAATTTATTCGGATCAATCAATGCTGAGACCAAGAGAGCCCTTTCTT 4229  
QY 4045 ggaatcaagtgtgactgtttggcccaatgattatcacggcatcaagaccgaagtgccc 4104  
Db 4230 GGATCAAGGTGGGATGTTGGCCACCAATGATTTTCACGGCATCAAGACCAAGGGTGCCC 4289  
QY 4105 gtcaagaagttctccagccttcacatctctcaagttatcaatcagtaagttatgagtgga 4164  
Db 4290 GTCAAGAGTTCTCCAGCCTCTACATCTCTCAGTTTATCATCATGATATGCTTGTGATGGGA 4349  
QY 4165 agaagtggaagactatcagagaaattccactggaaccttaatggtctctcttggaatg 4224  
Db 4350 AGAAGTGGCAGACTTATTCAGAGAAATTCACCTGGAACCTTAATGCTTCTTGGCAATG 4409  
QY 4225 tggattcatctgggataaaacaataattttaaaccctccaattatgtctgataatcc 4284  
Db 4410 TGGATTCATCTGGGATTAACACAAATATTTTAAACCTCCAAATATTTGCTCGATACATCC 4469  
QY 4285 gtctgaccaactatctatagatctgcagacacctcttcgcatggagttgagtgctg 4344  
Db 4470 GTTTCGACCCCACTCTTATATAGCATTCGCAACACTCTTGCAATGGAGTTGATGGGTTG 4529  
QY 4345 atttaaatgtgcaagcatgycatttgggaatggagagtaaaagcaatataagatgacaga 4404  
Db 4530 ATTTAAATAGTTTCAGACATGCCCATTGGGAATGAGAGTAAGCAATATCAGATGACAGAG 4589  
QY 4405 ttactgtctcatctacttaccataatgtttggcaacctggtctccttcaaaagctcgac 4464  
Db 4590 TTAAGCTTTCATCTTACTTATCCAAATATGTTTCCACCTGGCTCTCTTCAAAAAGCTCGAC 4649  
QY 4465 ttcaactccaaggaagaaatgaatgctgagacactcaggtgaataatccaagaagtggc 4524  
Db 4650 TTTCACCTCCAGGAGAGATGAATGCTTGAGACTCAGGAGATTAATCCAAAAGAGTGGC 4709  
QY 4525 tgcgaatggaacttcagaaagaaatgaagtcacaggaagtaactactcagggagtaaat 4584  
Db 4710 TGCAGTGGACTTCCAGAGACAAATGAAGTCCAGAGAGTAAGTCAAGGAGTAAAT 4769  
QY 4585 ctctgtctccagcatgtatgtgaagagttcctcaatcccaagcaataagatggccatc 4644  
Db 4770 CTCTGCTTACCAACATGATGTGAAGAGTCTCTCATCTCCACCACTCAAGATGGCCATC 4829  
QY 4645 aatgagactctcttttccgaagtgaagaagtttccagggaaatacaagactcct 4704  
Db 4830 AGTGGACTCTCTTTTTCAGAAATGCAAGTAAAGSTTTTTCAGGAAATCAAGACTCCT 4889  
QY 4705 tcaacactgtgtgaactctctagaccacgcttactgactcgtaactcctgaattcac 4764  
Db 4890 TCACACCTGTGGTGAAGCTCTTAGACCAACCGTTACTGACTCGCTACCTCGAATTCAC 4949  
QY 4765 ccacagagttgggtgacacagatgcccctgaagatggaagttctggctgcaagacag 4824  
Db 4950 CCCAGAGTTGGGTGACACAGATTTGCCCTGAGATGAGGTTCTGGCTGCGAGGACAG 5009  
QY 4825 acctactgactcagcgaggttc 4848  
Db 5010 ACCTTACTGAGGGTGGCCACTGC 5033

RESULT 11  
US-08-670-707A-38

Sequence 38, Application US/08670707A  
Patent No. 5859204  
GENERAL INFORMATION:  
APPLICANT: Lollar, John S.  
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Circle Suite 201  
CITY: Boulder  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/670,707A  
FILING DATE: 26-JUN-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US94/13200  
FILING DATE: 15-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/212,133  
FILING DATE: 11-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/864,004  
FILING DATE: 07-APR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Greenlee, Lorraine L.  
REGISTRATION NUMBER: 27,894  
REFERENCE/DOCKET NUMBER: 75-95F  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 303/499-8080  
TELEFAX: 303/499-8089  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4334 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: not relevant  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Factor VIII lacking B domain  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 3..4334  
US-08-670-707A-38  
Query Match 40.2%; Score 3189.8; DB 2; Length 4334;  
Best Local Similarity 83.5%; Pred. No. 0;  
Matches 3692; Conservative 0; Mismatches 637; Indels 90; Gaps 3;  
QY 420 atgcaaatagagctccacactgctcttctctgcttgcttgcttgatctgcttgaagtc 479  
Db 3 ATCCACTTGAAGCTCCACACTGTGCTTCTGTGTCTGTGCTGTGCACTCGGCTTTAGTGC 62  
QY 480 accagaagatactactgggtgagtggaactgcatggaactatatgcaaatga---t 536  
Db 63 ATCAGAGATATCTACCTGGGCGAGTGAAGTGTCTCTGGGACTACCGGCAAAAGTGAATC 122  
QY 537 ctctggtgactgctgtgtgaacgaagaatttcctcctgaagtgcaaaacttttcattc 596  
Db 123 CTCCTGAGCTGCACGTGACACACAGATTTCTGCTTACAGCGCAGAGACTCTTCCGTTG 182  
QY 597 aacactcagctggtgacaaaagactcgtgtgtagaattcaagggttaacctttcaac 656  
Db 183 GGCCTGCTAGCTCTGTATCAAAAAGACTGTGTCTGTAGAGTTACGGATCATCACTTTTCA 242

|    |      |        |                 |        |         |        |        |       |        |          |        |      |
|----|------|--------|-----------------|--------|---------|--------|--------|-------|--------|----------|--------|------|
| OY | 657  | atcgc  | taagccaaagccacc | ccctg  | atggtgc | ctc    | taagtc | ccct  | tacac  | cccaagct | gaagtc | 716  |
| OY | 658  | ttt    | ttt             | ttt    | ttt     | ttt    | ttt    | ttt   | ttt    | ttt      | ttt    | 717  |
| Db | 243  | gtttgc | acagccccc       | acagcc | acacat  | ggatgg | gggtct | ctg   | tggtct | ctac     | catcc  | 302  |
| OY | 717  | tatgt  | taagagtg        | tac    | tac     | tac    | taagaa | tggct | cccat  | ccct     | gtc    | 776  |
| Db | 303  | tacgac | ac              | cggtg  | ctt     | ac     | ctt    | ac    | ctt    | gaagaa   | cat    | 362  |
| OY | 777  | gtctg  | tgt             | ta     | ccctac  | ctg    | gaagct | ctc   | tg     | agagag   | ctg    | 836  |
| Db | 363  | gtcggg | ctc             | ttc    | ctt     | gt     | gaaat  | ctt   | cc     | gaaggg   | cgctga | 422  |
| OY | 837  | aggga  | aaagaaga        | atga   | tga     | atc    | ttcc   | ctc   | gtg    | tgaag    | ccata  | 896  |
| Db | 423  | aggga  | aaagaaga        | acat   | atga    | atc    | ttcc   | ctc   | gtg    | tgaag    | ccata  | 482  |
| OY | 897  | ctgaa  | agagaat         | atg    | tccaa   | tgc    | ctc    | tg    | ac     | ccac     | ctg    | 956  |
| Db | 483  | ctgaaa | agaaaat         | atg    | tccaa   | tgc    | ctc    | tg    | ac     | ccac     | ctg    | 542  |
| OY | 957  | cat    | tg              | agcc   | ctt     | g      | tg     | taaa  | agac   | tt       | gaat   | 1016 |
| Db | 543  | cacgt  | gcac            | ctc    | gtg     | taaa   | agcc   | ctt   | gaat   | gc       | ctg    | 602  |
| OY | 1017 | gaa    | aggaag          | ctc    | tg      | ccaa   | gaagaa | agac  | agac   | act      | tg     | 1076 |
| Db | 603  | gaa    | aggaag          | ctc    | tg      | ccaa   | gaagaa | agac  | agac   | act      | tg     | 662  |
| OY | 1077 | g      | at              | t      | tg      | a      | tg     | a     | a      | a        | a      | 1136 |
| Db | 663  | g      | at              | t      | tg      | a      | tg     | a     | a      | a        | a      | 722  |
| OY | 1137 | gat    | ctg             | ta     | ctg     | ta     | ctg    | ta    | ctg    | ta       | ctg    | 1196 |
| Db | 723  | gat    | ctg             | ta     | ctg     | ta     | ctg    | ta    | ctg    | ta       | ctg    | 782  |
| OY | 1197 | tct    | ctg             | cc     | ag      | gt     | ct     | gt    | tg     | ta       | tg     | 1256 |
| Db | 783  | tct    | ctg             | cc     | ag      | gt     | ct     | gt    | tg     | ta       | tg     | 842  |
| OY | 1257 | ggc    | acac            | cc     | ctg     | aag    | tg     | ca    | ct     | ca       | ta     | 1316 |
| Db | 843  | ggc    | acac            | cc     | ctg     | aag    | tg     | ca    | ct     | ca       | ta     | 902  |
| OY | 1317 | cat    | cg              | ca     | ag      | cg     | ct     | ct    | tg     | ga       | a      | 1376 |
| Db | 903  | cat    | cg              | ca     | ag      | cg     | ct     | ct    | tg     | ga       | a      | 962  |
| OY | 1377 | a      | tg              | gc     | ct      | tg     | gc     | ct    | tg     | gc       | ct     | 1436 |
| Db | 963  | a      | tg              | gc     | ct      | tg     | gc     | ct    | tg     | gc       | ct     | 1022 |
| OY | 1437 | ga     | ag              | ct     | a       | tg     | ca     | a     | g      | ct       | g      | 1496 |
| Db | 1023 | ga     | ag              | ct     | a       | tg     | ca     | a     | g      | ct       | g      | 1082 |
| OY | 1497 | ga     | aga             | ag     | cg      | ga     | ag     | ct    | a      | tg       | a      | 1556 |
| Db | 1083 | ga     | aga             | ag     | cg      | ga     | ag     | ct    | a      | tg       | a      | 1139 |
| OY | 1557 | gat    | ga              | ga     | aa      | ac     | ct     | ct    | cc     | ct       | ta     | 1616 |
| Db | 1140 | gat    | ga              | ga     | aa      | ac     | ct     | ct    | cc     | ct       | ta     | 1199 |
| OY | 1617 | act    | gt              | ga     | t       | ca     | t      | a     | c      | a        | t      | 1676 |
| Db | 1200 | act    | gt              | ga     | t       | ca     | t      | a     | c      | a        | t      | 1259 |
| OY | 1677 | g      | o               | c      | c       | c      | g      | a     | t      | a        | g      | 1736 |
| Db | 1260 | g      | o               | c      | c       | c      | g      | a     | t      | a        | g      | 1319 |

|    |      |   |      |
|----|------|---|------|
| OY | 1737 | aggaagacacaaaaaaagccgattcttggcgttaaacagatggaacctttaagccctgtga   | 1796 |
| Db | 1320 | AGGAATAATCAAAAAAGCTCGATTTCGTGCTTACACGGATGTAACTTTAAGACTGTAAA     | 1379 |
| OY | 1797 | gcataccaagaatcaagaactcttggaaccttcctaagggaagtltgagacaca          | 1856 |
| Db | 1380 | GCTATTCCGATGATCAGAATCCGTGGACCCTTTACTTTATGGAGAAGTTGGACACA        | 1439 |
| OY | 1857 | cgttgttatatatltaaagaatlcaagcaagcacacataaatctaccctcaagaaic       | 1916 |
| Db | 1440 | CTTTTGATTATTATTAAAGTAAGGAGCGCACCATTAATACATCTACCTCATGGAATC       | 1499 |
| OY | 1917 | actgatgccgtccccttgtaattcaagagatlaaccaaaagctgtaaaaacattigaagat   | 1976 |
| Db | 1500 | ACTGATGTCAACGCTTTTGACACCGAGAGACTCTMAAAGTTGGAAACATTTGAANAAC      | 1559 |
| OY | 1977 | tttccaattctgcagaganaaatatccaatalaaatgtagacygactgtlagagaatgg     | 2036 |
| Db | 1560 | ATGCCAATTCCTGCAGAGAGACTTTTCAMATTAATAATGACAGTGCATGTGGAAATGGG     | 1619 |
| OY | 2037 | ccaactaaatcaagatctcgctggctgcagcccgatatctactagttctgataatgag      | 2096 |
| Db | 1620 | CCAACCAAGTCCATCTCTCGGTGCTGACCCGCTACTACTCAGCTCCATTATCTAAG        | 1679 |
| OY | 2097 | agagattcagcttlccaagcatcaattggccctctccatctgtctacaaaalctytagat    | 2156 |
| Db | 1680 | AAGATCTGGCTTGGGAGACTATTTGGCCCCCTCTCATCTGCTAACAAAGAACTGTATAC     | 1739 |
| OY | 2157 | caaagagaanaaacaagaataatgtagaagaagaagaatgatactcttctgtattat       | 2216 |
| Db | 1740 | CAAAGAGAAACACGATGATGTGACACAGAAACGTATCCTGTCTGTATTCAT             | 1799 |
| OY | 2217 | gagaccgaagctgtgtaeccacagagaatatacaagcttctccccaatccagcttga       | 2276 |
| Db | 1800 | GAGAAATCAAGSCTGTGATCTCGAGAGAAATTTACAGCGTTCTCTCCCAATCCGATGCA     | 1859 |
| OY | 2277 | gtgcagcttgaggaatccagaagttccaaagcctccacaatcaycacagaatcaygctat    | 2336 |
| Db | 1860 | TTCACGCCCCAGATCCACAGATTCCAAAGCTTTCAACATCATCACAGATCATMTGGCTAT    | 1919 |
| OY | 2337 | gttttbatgtttgacagttgttcaagtttbtgcaaggatggacactgtgatactta        | 2396 |
| Db | 1920 | GTTTTTGATTAACCTTGACAGCTGTGGCTTTGTTTTCACAGAGTGGCATCTGTACTTCTA    | 1979 |
| OY | 2397 | agcaatlgagacacagactgtaactctctctgtctctctctcgyaatalaactctcaaac    | 2456 |
| Db | 1980 | AGTGTGGAGACACAGAGSGACTTCCTCCGTCTTCTCTGGCTACACCTTCAAAAC          | 2039 |
| OY | 2457 | aaaatgtctatgtagaacacaactacacocctatccccattcgagggaactgtctcgt      | 2516 |
| Db | 2040 | AAAATGGCTATGAAGACACACTCAACCTGTTCCCTTCCAGAGAAAAGSTCTTTCATG       | 2099 |
| OY | 2517 | tcgaatggaaaaaccagaatctatggaatctctgggtgtgcacaactcagaactltgagaaca | 2576 |
| Db | 2100 | TCAATGGA AAAACCCAGSTCTCTGGGTCTTAGGGTGCACAACTCAGACTTGGGAAACAG    | 2159 |
| OY | 2577 | ggcatgaccgccttaactcgaagagttcttagtgtgacaagaacaacgtygatataagag    | 2636 |
| Db | 2160 | GGGATGACAGCCTTAACGAAGGGTGTATGTATGTGAACGGAGACATTTGGTATTTATATAC   | 2219 |
| OY | 2637 | gacagttatgaagatatttcaagataactgtctgagtaaaaaacaatgaccaatlaacaaga  | 2696 |
| Db | 2220 | AACACTTATGAAGTATTCCACAGCCTTCTTGCTAGTGGAAAGATGTCAATTGAACCCAGA    | 2779 |
| OY | 2697 | agcttctccagaaatlcaagaacacccatgacatlagycaaagooaaltaatgcaaccca    | 2756 |
| Db | 2280 | -----   | 2279 |
| OY | 2757 | ccagttcttgaagcccatcaacgggaataactcgtactactcttcaugtccagataagag    | 2816 |
| Db | 2280 | -----GACATTAAGGCTTTCCTTACTTTTTCAGCCGGAGGAATAC                   | 2315 |
| OY | 2817 | gaaattbaactatgataacacatacagtttgaatgaagaagaagaatttbaacattat      | 2876 |

Db 2316 AAAATGACTATGATGATATCTTCAACTGAACGAAGGAGAGAATTTTGACATTAC 2375  
QY 2877 gatagatgataaatacagagcccccgaagccttcaaaaaaacaacgaacataattatc 2936  
Db 2376 GGTAGAGATGAAATGACAGACCTCGCACTTTCAGAAAGAACCCGACATTATTTAT 2435  
QY 2937 gctcaatgagagagcctctggtatataaggaatgagtagctcccaatgctctaagaac 2996  
Db 2436 GCTCGGGGAGAGAGCTCTGGGATTTACGGGATGAGCCGATCCCGCGGCTTAAGAAC 2435  
QY 2997 agggcctaagatggcaatgctccatgaatgaagaagtgtttccagaattactgat 3056  
Db 2496 AGGGCTCAGAACGGAGAGGTCTCGGTTCAAGAGGTGCTTCCGGGATTTGCTGAC 2555  
QY 3057 ggcctcttactcagcccttataccgtgagaaactaaatgaacatttgggaactctggg 3116  
Db 2556 GGCTCCTTCAGAGCGCGCTCTACCGGGGMACTCAACAACTTGGGGCTTGGGA 2615  
QY 3117 ccatataaagacagaagttgaagataatacatatgtaacttcaagaatacagcctc 3176  
Db 2616 CCTTACATCAGAGCGGAGTTTGAAGACATCATGTTCAAAACCCAGCGCT 2675  
QY 3177 cgtccattctctctatctcagccttattctataggaagaatcagaggaagca 3236  
Db 2676 CGTCCATTCTCTCTACGAGCCTTATTCTTATCCGGATGATCAGAGCAAGGGCA 2735  
QY 3237 gaaccttgaaaaaactctgtcaagcctaatagaacaaactactcttggaaagtca 3296  
Db 2736 GAACCTTGACACAACTTCGACGCAAAATGAACCCAGAACTTCTTGGAAAGTCCAG 2795  
QY 3297 catcatatgcaaccactaaagatgatttgaactgaagcctgggtttatctctgat 3356  
Db 2796 CATACATGAGACCCACAGAAAGAGAGATTGACTGCAAAAGCCTGGCTACTTTCTGAT 2855  
QY 3357 gttgaacctggaaaaaatgctgactcagagcctgaatggacccctctgcttgcacact 3416  
Db 2856 GTTACCTGGAAGAAAGATGTCACACTAGGCTTGATCGCCCTTCGATCTGCGCGCC 2915  
QY 3417 aacacattgaacctgctctatggagaagaatgagcagtaagaagatttgcctgttc 3476  
Db 2916 AACGCTGAAACGCTGCTCACGGTAGACAAAGTACCGCTGCAAGAAATTCCTGTTTTTC 2975  
QY 3477 accactctgatgagacaaagaactgtgtacttcaactgtaaaatagagaagaacactgag 3536  
Db 2976 ACTATTTTGTATGAGAAAGAGCTGTACTCTCATGAAATGTGGAAAGAACTGCCGG 3035  
QY 3537 gctccctgcaaatccagatggaagatcccaactttaaagaaatatacgcttccatgca 3596  
Db 3036 GCCCCTGCAACGTCAGATGAGAGACCCCACTGAAAGAAACATTCGCTTCATGCA 3095  
QY 3597 ataataatgctcaataatagatacactcgtgcttagtaatgctcagatcaagaatc 3656  
Db 3096 ATCATATGGCTATGTGATGATACCTCCGCTTGTAGTAATGGCTCAAGAAAGATC 3155  
QY 3657 cgaatgatactgctcagcagcagcagcaatgaaacatccattcatctatcttaagcga 3716  
Db 3156 CGATGGTATCTGCTCAGACATGGGCGACAAATATTCATTGATTCATTTTAGCGGA 3215  
QY 3717 catgtgtcactgtaagaaaaaagaagagatataaaatggcactgtatacctctatcca 3776  
Db 3216 CAGGTGTACGTACGAGAAAGAGAGAGATATTAATGCGCGGTACATCTCTATCCG 3275  
QY 3777 ggtgtttttagaagaaggaatgttaacatccaaagcttggaatttggcgggtggaaatgc 3836  
Db 3276 GGTGTCTTTAGAGAGTGAATCTACCGCTCAAAAGTTTGAATTTGGCAATAGATTCG 3335  
QY 3837 cttaatggcagcactacatgctcgtgatgagcacactttctcgtgtatacgaacaatg 3896  
Db 3336 CTGATTTGGCAGACCTCGCAAGCTGGGATGAGCAGACTTCTCGGTGATACGCAAGAG 3395  
QY 3897 tctgaactccctgggaatgctcttgcagacaattagagatttcaagatacagcttca 3956  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 3396 TGTGAGCTCCACTGGGAATGCTTCTGAGACGATTAAGATTTTCAGATCACAGCTTCA 3455  
QY 3957 ggaacaatatgagcagctggccccaagaactggccagacttcatcttccgatatcaaat 4016  
Db 3456 GGACGATATGACAGTGGGCCCCAAAGCTGGCGAGCTTCAATTTTCGGGTCAATCAAT 3515  
QY 4017 gcttggagcaacgaagagcccttctcttggatcaaggttgatcgtgttggcaaccaatgat 4076  
Db 3516 GCCTGAGACACCAAGATCCCGACTCTCGTATGATCAAGGTGATCTGTTGGACCAATGATC 3575  
QY 4077 attcaagcatalcaagaccgaaggtgcccgtcaagaagttcttccagcctcatcatctcag 4136  
Db 3576 ATTCCAGGATCATGACCAAGAGGTGCCGTCAGAAATTTTCCAGCCCTCATCTCCAG 3635  
QY 4137 ttatcatatgataatcttcttgaatggagaagatggcagagacttctgaagaattccact 4196  
Db 3636 TTTATCATCATGTACAGCTTGTGAGGAGAACTGGCAGAGATTTCCAGAGGAATTCACG 3695  
QY 4197 ggaaccttaatgctctctcttggcaatgltgaltcatcttggaataaacaacataattt 4256  
Db 3696 GGCACCTTAATGCTCTTCTTGGCAATGTGACCCATCTGGGATTAACACAAATATTTT 3755  
QY 4257 aaccttcaattatgtctcgaatatacatcgttgcaccaactcatataagcattcagc 4316  
Db 3756 AACCTCCGATTTGGTCCGTACATCCGTTTGCACCCAAACATTTACAGATCCGAGC 3815  
QY 4317 accttcgcataatgagatgagtgctgtatataatgttgcagcagatgcatgggaatg 4376  
Db 3816 ACTCTTGCAATGAGTGTGATGGGCTGTGATTTAAACAGTTTGCACACAGCCCTTGGAATG 3875  
QY 4377 gagaataaacaatatcagatgacagattactgtcttcatctcattcaacataatgtt 4436  
Db 3876 CAGATTAAGAGATATGAGATCAGATCAGAGGCTCTCCACCTCAGCAATATATTTT 3935  
QY 4437 gcaacctgtctctctcaaaaagctcgacttcaacctcaaggaaggaatgctctggaga 4496  
Db 3936 GCCACCTGTCTCTCTTCAAGCCCGACTTCACTCCAGGGGCGAGGAATGCTTGCGGA 3995  
QY 4497 cctcaggtgaataatccaaaagatggctgcaatggaacttccagaagaataataaagtc 4556  
Db 3996 CCCCgggtgAGACGACGAGAGAGAGTGTGAGGTGGAACCTGCGAAGAGACGTTAAAGTTC 4055  
QY 4557 acaagataacttactcagaggaagtaaaatctctgtcttccacagcatgtatgtgaagagtc 4616  
Db 4056 ACAGGCAATCACACCCAGGGCGTGAATCCCTGCTCAGACAGATGATGTGAAGAGATTTC 4115  
QY 4617 ctcatctccagcagatcaagaatgycalcagltggaactctcttcttcagaaatgycaaagta 4676  
Db 4116 CTGCTGTCAAGTATGTCAGAGCGCGCGCTGGACCTGTTCTTTCAGAGCGGCCACAG 4175  
QY 4677 aaggttttcaaggaataatcaagactccttcaacactgtgttgaactctctagaccacg 4736  
Db 4176 AAGGTTTTTCAAGGCAATCAGGATCTCTCAACCCCGTGTGAAGCTGTGGACCCCGC 4235  
QY 4737 ttaactgcctcgaactctgaattcaaccccaagatgggttgggttgcacaagatgcccag 4796  
Db 4236 CTGTTACGGGCTACTCTGAGAGATCCACCACAGACTGGGCGCAGCATCGCCCTAGG 4295  
QY 4797 atggaagttctggctgagcagcaggaacctactga 4835  
Db 4296 CTCGAGGTTCTAGAGTGTGAGGACACAGGATCTCTACTGA 4334

RESULT 12  
US-09-601-38  
; Sequence 38, Application US/09037601  
; Patent No. 6180371  
; GENERAL INFORMATION:  
; APPLICANT: Lollar, John S.  
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.





[illegible]

|    |      |  |      |
|----|------|--|------|
| Db | 2436 | GCCTGCGTGGAGACAGCTCTGGATTAAGGATGAGCGAAATCCCCGGGCGCTAAGAAAC     | 2495 |
| Qy | 2997 | agggtcagatgtgcagatgtccctcagttcaagaagtgttttccagaatttactgat      | 3055 |
| Db | 2496 | AGGGCTCAGAAACGGAGAGTCCCTGGTTTAAAGAGGTGGTCTTCCGGAAATTTGGCTGAC   | 2555 |
| Qy | 3057 | gacctccttactcagaccttaccgtyggaactaaatgaacatttggaactccctggag     | 3111 |
| Db | 2556 | GGCTCCTTCACGCAACCGTCGTACCGGGGGAACTCAACAACACTTGGGGCTCTTGGGA     | 2611 |
| Qy | 3117 | ccataataagagagaagtltgaagataatcatlytlaacttccagaatcgaagcctct     | 3176 |
| Db | 2616 | CCCTACATCAAGAGCGGAAGTTGAAGACACATCATGTGTACTTTCAAAAACCGGGTCT     | 2675 |
| Qy | 3177 | ggtccctatctccctctatcttctagccttatttcttataagaagaatacagaagcagca   | 3233 |
| Db | 2676 | CGTCCCTAATTCCTTCTACTCAGCGCTTATTTCTTATCCGATGATCAGGAAGGGGCA      | 2733 |
| Qy | 3237 | gaacctagaanaaacttltgaagccctaaagaaccaaaacttacttltggaagtgtgca    | 3296 |
| Db | 2736 | GAACTGCACACAACTTCGTCTCAGCCCAATGAAACAGACTTACTTGTGGAAAGTGCA      | 2795 |
| Qy | 3297 | catcataatgcaaccactaaagaatltgtactgtcagaagcctggacttattcttctgat   | 3355 |
| Db | 2796 | CATCAGATGGCACCACAGAAAGAGATTGTACTGCAAAAGCCGGGCTACTTTTGTAT       | 2855 |
| Qy | 3357 | gttgaactgtgaaaaagatgtgtcactcagcgcgtatgtgaaccttctgtctgcacact    | 3416 |
| Db | 2856 | GTTCACCTGGAAAAAGATGTGCATCAGGCTTATGTGGCCCTCTGTGATCTGCGCGCC      | 2915 |
| Qy | 3417 | aacacatgaacctgtctatctggaagaagaatgaagaacgaatttgcctgttttct       | 3476 |
| Db | 2916 | AACACCTGGAACGCTGCTCAGCGTAGACAACTGACCGTGCAAAATTTGCTGTGTTTT      | 2975 |
| Qy | 3477 | accatcttltgaatgaaccaaagctgttlaacttaactgaaataatgtgaaagaacatgcag | 3536 |
| Db | 2976 | ACTATTTTGTGATGAGACAAAGAGCTGTACTTCACTGAAAATGTGGAAAGAAAGTCCGG    | 3035 |
| Qy | 3537 | gtctccctgcacatctcagaatgtgaagatccacattttaaagaataatctgtcttccatga | 3596 |
| Db | 3036 | GGCCCCGTGCCACCTGCAGATGGAGGAGCCACCTCTAAGAAAGAACTATTCGTTCCATGCA  | 3095 |
| Qy | 3597 | atgaatgtcacaataatgagatacactaccgtgttgatgaatgtgtcagaatacgaatgatt | 3656 |
| Db | 3096 | ATCAATGGCTATGTGATGATACACTCTCGGTTAGTATGGCTCAGATTAAGAGATC        | 3155 |
| Qy | 3657 | cgatgtatctgtctcagaatgcagcagcagaatgaaataatccatcttacttcaattgtga  | 3716 |
| Db | 3156 | CGATGTATCTGCTCACCAATGGGAGCAATGAATAATTCATTCATTATTTTACGGGA       | 3215 |
| Qy | 3717 | catgtgttccactgtagaanaaaaaagaagaglataaatgycacgttgcacatctcatcca  | 3776 |
| Db | 3216 | CACGTCTTCAGTGTACGAAAAAGAGAGATTAATAATGGCCGTGACAAATCTATACCG      | 3275 |
| Qy | 3777 | gggttctttagaagaatgtgaatgttaccatccaaagcctggaatttgggggtgaagtc    | 3836 |
| Db | 3276 | GGGTCTTTGAGAGAGGAAATGTCTACCTCCAAAAGTTGGAAATTTGGCAATTAAGATC     | 3335 |
| Qy | 3837 | cttatttgcagacatctacaatgtcttggaatgaagcacttcttctgttgaagcaataag   | 3896 |
| Db | 3336 | CTGATTTGGCGAGCAACGTGGAAGTGGGATGAGACGACACTTTCCTGTGTACGCAAGAG    | 3395 |
| Qy | 3897 | ttgtcagaactccctgggaatgtgtctctgtgcacatataagaatttcaagtatacagctca | 3956 |
| Db | 3396 | TGTACGCTCCACTGTGGGAAGGCTTCTGTGACGATATAGAGATTTTCAATCAACACTTCA   | 3455 |
| Qy | 3957 | ggacaataatggaagaatgtggccccaagcctgtgcagactcatattccggatccaataat  | 4016 |
| Db | 3456 | GGACAGTATGACAGTGGGGCCCAAAAGCTGGCCAGACTTCAATTAATTCGGATCAATCAT   | 3515 |
| Qy | 4017 | ggcttgagacccaagagaccttctcttgatcaagaatgtatgtctgtgcaccaatgtt     | 4076 |



|    |      |  |      |
|----|------|--|------|
| Db | 3516 | GCCTGAGACACCAAGATCCCACTCCTTGATCATAAGGTGATCTGTTGGACCAATGATC       | 3575 |
| Qy | 4077 | attcaacgcatacaagaccgaaggtgcgccgtccagaagttctccagcctctacatctccag   | 4136 |
| Db | 3576 | ATTTCAGGAGATCATGACCCAGGGGTGCCGTCAGAAAGTTTTCACACTCTACATCTCCAG     | 3639 |
| Qy | 4137 | ttatcatcatgtatgactctgtgtggaagaagtggaagactcttgaagaatctcaact       | 4196 |
| Db | 3636 | TTTATCATATGTACAGTCTTTGACGGGAGGAACTGGCAGAGTTTCCAGGGGAATTTCCAG     | 3695 |
| Qy | 4197 | ggaaccttaatgtctctctcttgcgaatgttgattcatctcgtggaataaacaatatltt     | 4256 |
| Db | 3696 | GGCACTTATATGCTCTTCTTTGGCAATGTGAGCAGCATCTGGAGTTAAACAATATTTTTT     | 3755 |
| Qy | 4257 | aacctccaatatatgtctcgtatataatccgtttgcaccaactataatgacttcgaac       | 4316 |
| Db | 3756 | AACCTTCCCAATTTGTGGCTCGGTACATCCGTTTGGACCAACACTTTCACAGATCCGAGC     | 3815 |
| Qy | 4317 | actcttcgatbtgaagttgaaggctctgtattaaatgattgcgaatgcacatgcattggagatg | 4376 |
| Db | 3816 | ACTCTTCGATGAGATTGATGGGCTGTGANTTAAACAGTTTGCACAGTCCCTCTGGAGTG      | 3875 |
| Qy | 4377 | gagagtaaaagcaataatcagatgacagatattactgcttcatcctacttaccatatgtt     | 4436 |
| Db | 3876 | CAGAATAAAGCGATATCAGACATCACAAGATCAGCGCCCTCTCCACCTTAAGCAATATATTT   | 3935 |
| Qy | 4437 | gccacctgtctctccttcaaaagctcgaactcactccccaaggagagtaatgtcctgaga     | 4496 |
| Db | 3936 | GCCACTGGTCTCTTCACAAAGCCCACTTCACCTCCAGGGGCGGAGCAAGTCCCTGGCGA      | 3995 |
| Qy | 4497 | cctcagtgtaataatccaaaagggtgcctcgaatggaactccagaagaacataaaagtc      | 4556 |
| Db | 3996 | CCCCGGGTACGACGCCAGAGAGTGGCTCAGGTGGAACCTGCAGAGACGCTAAAGTTC        | 4055 |
| Qy | 4557 | acaaagtaactactcaagagagtaaaatctctctcttccaccagcatgtatgtaaagagttc   | 4616 |
| Db | 4056 | ACAGGATACACCAACCGAGGGCGGTGAAGTCCCTGCTCAGCAGCATGTATGTAAAGAGTTC    | 4115 |
| Qy | 4617 | ctcaatccagcagtcataagaatgtgcataagtggaactctcttcttcagaatbgtcaaatga  | 4676 |
| Db | 4116 | CTCGTGTCAGATGATCAGAGCGCGCCGCTGGAGACCTGTTCTTCTTCAGGACGGCCACACG    | 4175 |
| Qy | 4677 | aaggttttccagggaatacaagaagctctcttcaacacgttgatgaactcttagaccacgcg   | 4736 |
| Db | 4176 | AAGGTTTTTCAGGGCAATCAGGAGCTCTTCACCCCGGTGTGAACGCTGTGGACCCCCCG      | 4235 |
| Qy | 4737 | ttactgactcgtactaccttcgaattcaaccaccagaagttgtgacacgaattgtccctgag   | 4796 |
| Db | 4236 | CTGTTACAGCGGTACTCTGAGAGATCCACACCACGACAGTGTGGGGCAGACATCGCCTTGAG   | 4295 |
| Qy | 4797 | atggaagttctggtcgtcgagagcacagaagactctaetga                        | 4835 |
| Db | 4296 | CTCGAGGTTCTAGAGATGTGTGGCGACAGAGATCTTACGGA                        | 4334 |

RESULT 13  
US-08-893-327-15  
Sequence 15, Application US/08893327  
Patent No. 6020192  
GENERAL INFORMATION:  
APPLICANT: Zolotukhin, Sergei  
APPLICANT: Hauswirth, William W.  
APPLICANT: Muzyczka, Nicholas  
TITLE OF INVENTION: Humanized Green Fluorescent Protein  
TITLE OF INVENTION: Genes and Methods  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P. O. Box 4433  
CITY: Houston  
STATE: TX  
COUNTRY: USA  
ZIP: 77210-4433

```

1 COMPUTER READABLE FORM:
2
3 MEDIUM TYPE: Floppy disk
4
5 COMPUTER: IBM PC compatible
6
7 OPERATING SYSTEM: PC-DOS/MS-DOS
8
9 SOFTWARE: PatentIn Release #1.0, Version #1.30
10
11 CURRENT APPLICATION DATA:
12
13 APPLICATION NUMBER: US/08/893,327
14
15 FILING DATE:
16
17 CLASSIFICATION: 514
18
19 PRIOR APPLICATION DATA:
20
21 APPLICATION NUMBER: US 08/588,201
22
23 FILING DATE: 18-JAN-1996
24
25 ATTORNEY/AGENT INFORMATION:
26
27 NAME: Kitchell, Barbara S.
28
29 REGISTRATION NUMBER: 33,928
30
31 REFERENCE/DOCKET NUMBER: UFLA:062\KIT
32
33 TELECOMMUNICATION INFORMATION:
34
35 TELEPHONE: (512) 418-3000
36
37 TELEFAX: (713) 789-2679
38
39 INFORMATION FOR SEQ ID NO: 15:
40
41 SEQUENCE CHARACTERISTICS:
42
43 LENGTH: 6253 base pairs
44
45 TYPE: nucleic acid
46
47 STRANDEDNESS: single
48
49 TOPOLOGY: linear
50
51 FEATURE:
52
53 NAME/KEY: CDS
54
55 LOCATION: 988..1701
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
```

[illegible]

Dh 3787 CTGTCCGACCCCTCCGCTTACCGGATACCTGTCCGCCCTTCTCCCTCGGAAAGCGTGG 3846  
OY 5516 cgtcttccaatgctcagcgtgtagtatactcagtcggtgtagtgcgttcgccaagc 5575  
Dh 3847 CGCTTTCTCAATGCTCAACGCTGTAGTATCTCACTTCGGTGTAGTGTCTTCCCAAGC 3906  
OY 5576 tgggctgtgtgcaagacccccctcagccgacgcgtgcgcttaccgtaatac 5635  
Dh 3907 TGGGCTGTGTGACGAACCCCGCTCAGCCGACCGGTGGCTTATCCGGTAATCATC 3966  
OY 5636 gtctttagtccaacccggtlaagacagactatcgccaatgycagcgaactgtaaca 5685  
Dh 3967 GTCTTGAGTCCAAACCCGTAAGACACGACTTATCGCCATGGCAGCCACTGTATAA 4026  
OY 5696 ggtatcaagagcgaagtatgtagcggtgtacagaagcttcaagtgtgggctaact 5755  
Dh 4027 GGATTAGAGAGCGAGGATGTAGGCGGTGTACAGAGTTCTTAAGGGTGGCTTAATCT 4086  
OY 5756 acggtacaactagaagagagatattgtatctgcgtctgtgaagccaatlaactcg 5815  
Dh 4087 ACGGCTACACTAGAGAGAGATATTGGTATCTCGCTCTGTCTGAAGCCATTAACCTTCG 4146  
OY 5816 gaaagaagtgtgtagctcttgatcgcgcaaaacaacccgcgtgtgtagcggtttt 5875  
Dh 4147 GAAAAAGAGTGTGAGCTTTGATCCGGCAAAACACCGCTGTAGCGGTGTTT 4206  
OY 5876 ttofttcaagcagaattagcgcgagaaaaaaggatctcaagaagatcccttgaact 5935  
Dh 4207 TTGTTTCAGACGACCAATTAACGCGCAAAAAAGATCTCAAGAAAGATCTTTGATCT 4266  
OY 5936 ttcttaaggggtcgcagcgtcgaatgagcaagaaactcagctlaaggagattgtlca 5995  
Dh 4267 TTTCTACGGGGTCTGAGCGTCAGTGAACGAAACACACGTTAAGGATTTTGGTCAATGA 4336  
OY 5996 gattatacaaaaggatctcgaactagactcctttaaataaagaagtttaatacaa 6055  
Dh 4327 GATTATCAAAAAAGATCTTCACTAGATCCCTTTAAATTAATAATGAAGTTTAAATCAA 4386  
OY 6056 tctaaatataatgagtaactgtgtcagacttgaacagttgaactgaaggcac 6115  
Dh 4387 TCTAAATATATATGATTAACCTTGTGTGACAGTTACCAATGCTTATATGATGAGGCAC 4446  
OY 6116 ctatctcagcagatcgtctatcttcgtlcatcatcagttgctgcgactcccgctgtlaga 6175  
Dh 4447 CTATCTCAGGAGATGCTGCTATTTGCTTCATCATCAGTATGCTGACTCCGCTCGCTGATA 4506  
OY 6176 taactacgatacgggagggtctaccatctggccccagtgctgcgaatgatalacgcgagacc 6235  
Dh 4507 TTAATAGATACGGGAGGGGCTTACATCTGGCCCCAGTGTGCAATGATACCGCGAGACC 4566  
OY 6236 caggtcacacgggtccagatttaccagaataaaccagccagccggaaaggccgaagcca 6285  
Dh 4567 CACGCTACCGGGCTCCAGATTATACGAATAAACAACGACGCGGAAGGGCCAGAGCGCA 4636  
OY 6296 gaattggtcctcgaacttalcgcgtccatcagctcagctatlaattgtgtgcgggaacta 6355  
Dh 4627 GAATGTGCTCGCAACTTATCCGCCCTCCATCCAGTCTATTAATGTGTGCGGGGAAGCTA 4686  
OY 6356 gagtgaagtgtgcgaactaaatagtttgcgcaagctgtgtgcacattgtctacaggatcg 6415  
Dh 4687 GAGTAATAGTTCGCCAGTTAATAGTTGGCAACGTTGTTGCCATGTGTAACAGCATACG 4746  
OY 6416 tgtgtgaactgctgtgttggtagtgcgttcacatcagctcgcgtgtcccaacgaactcaagc 6475  
Dh 4747 TGGTGTACAGCTGCTGCTTTGGTATGGCTTCATTCACGCTCGGTTCCCAAGATCAAGGC 4806  
OY 6476 gagttaactgatcccccatgtgtgcgaaaaaagcgttagctccttcgtctcgcgactcg 6535  
Dh 4807 GAGTTACATGATCCGCCCATGTGTGCAAAAAAGCGTTAGTCTCTTGGTCTCCGATCG 4866  
OY 6536 ttgtcagaagaatagttggccgagtgltatcaactcagtgttatggcagcactgtcataatt 6595  
Dh 4867 TTGTCAAGAAAGTAAGTTGGCGCACTGTATCATCATGATGTTATGGCAGCACTCATTAAT 4926  
OY 6596 ccttactgtcatgcatccgtaagatgcttcttctgtgactggtgagtactcaaccaagt 6655  
Dh 4927 CTCTTACTGATATCCATCCGTAAGATGCTTTCTGTGACTGGGTAGTACTCAACCAAGT 4986  
OY 6656 catcttgagaatagtgtagtcgagcgaagtgctctcttcgccccggtlcaatacgggata 6715  
Dh 4987 CATCTGGAATAGTATGATGGCGACCGAAGTCTCTGTCGCCGCGTCAATACGGGATA 5046  
OY 6716 ataccggtccacatagaagaactttaaagtgtcctcatcatcttggaacagcttctcgggc 6775  
Dh 5047 ATACGGGCCCATAGCAAGAACTTTAAAGTGCATCATGTGAAAAAGTTCCTCGGGGC 5106  
OY 6776 gaaacctcaaggatcttaccgtgtgtgagaatccagttcagatgtaacccactgttcac 6835  
Dh 5107 GAATACTCTCAAGATCTTACCGCTGTGAGATCCAGTTCATTAACCCACTGCTGCAC 5166  
OY 6836 ccaactgtatctcagcatcttcttacttccacgaagcgttctggttgagcaaaaaagaa 6895  
Dh 5167 CCAACTGATCTTCAGCATCTTTACTTTACACGACGTTTCTGGGTGAGCAAAAAACAGAA 5226  
OY 6896 ggcataatgcgcgaaaaaaggaatagggcgacacgggaatgttgataactactc 6955  
Dh 5227 GGCATAATGCCGCAAAAAAGGAATAAGGCGCACCGGAAATGTTGAATACTCATACTCT 5286  
OY 6956 tccctttcaaatattatgaagcatltaacagggttcaagggttctgtcattgaagcgaatcat 7015  
Dh 5287 TCCCTTTTCATATATTTAGAGCATTTATATCAGGTTTATGTCTATGACGGAATCATAT 5346  
OY 7016 ttgaatgtattagaataataaacaataggggttccgcgcacattcccccgaatgtgc 7075  
Dh 5347 TTGAATGATTTAGAAAAATTAACAATAAGGGGTTCGCGCACATTTCCCGAAAGTGC 5406  
OY 7076 cactgaagcttaagaaacattatatacagatlaactataaataataggcgatca 7135  
Dh 5407 CACCTGACGCTTAAGAAACCAATTAATCATGACATTAACCTAATAAATGCGGTATCA 5466  
OY 7136 cgaagcccttgcgtcgcgcggttgcgtgtagtgcggttgaacactcgtacacatgcgc 7195  
Dh 5467 CGAGGCCCTTTCGCTGCGCGGTTCCGCTGATGACGGTGAAACCTGTGACACTGTGACG 5526  
OY 7196 tcccggagcgtgtacagactgtctgtlaagcgaatgcgcggagcagaacaagcccgtag 7255  
Dh 5527 TCCCGAGACGGTCAACAGCTGTCTGAAGCGATGCGCGGAGCAGAACCCGCTCAGG 5586  
OY 7256 ggcgcgtcagcgggtgtgtgcgcggtgtgcgggtgcgttactatgcgcacagacaga 7315  
Dh 5587 GCGGCTCAGCGGGGTGTGGCGGGTGTGCGGCTGCTTAATGATGCGCATCAGAGCAGA 5646  
OY 7316 ttgtactgagatgtcacacatatggtgtgtgaataacgcgcacgaatgtgtgaaggaatat 7375  
Dh 5647 TTGTACTGAGAGTCCACCATATGCGGTGTGAATAACCGCACAGATGCTAAGGGAATAT 5706  
OY 7376 accgcatacgaagaaatgtlaaacgttlaatatctgttlaaactcgcgttlaaattctgtla 7435  
Dh 5707 ACCGATCAGGAATGTGAACGTTAATATTGTTTAAATTCGCGTTAATTTTGTGTA 5766  
OY 7436 aatcagctcaattttttaaaccataatagccgaataatgcgcaaaaatccctataataaaga 7495  
Dh 5767 AATCAGCTCAATTTTAAACCAATAGCGCAAAATGCGCAAAATCCCTATTAATCAAAAGA 5826  
OY 7496 atagaccgagataaggtgtgtgtgtgttccagtttggaaacaaagagtcacatcaagaag 7555  
Dh 5827 ATAGACCGAGATAGGTTGAGTGTGTTCAGTTTGGAAACAAAGTCCACTATTAAGAA 5886  
OY 7556 cgttgactccaacgtlcaaaagggcgaanaacgltalacaaagcgaatggccaactacgtlga 7615  
Dh 5887 CGTGTACTCAACAGCTCAAAAGGGCAAAACCGTCTATCAGGGCATGGCCACTACGTGA 5946  
OY 7616 accatcaacctaatcaagtttttggggtgcgaagtgccgyltaagacactaatcggaaacc 7675  
Dh 5947 ACCATCACCCCTAAATCAAGTTTGGGTGTGAGGTGCGGTAACCAATCAATCGGAACCC 6006

QY 7676 taaaggagcccccgatttagagcttgaacgggaaagccgcgcgaactgtgcgagaaagga 7735  
| | | | |  
Db 6007 TAAAGGAGCCCCGATTTAGAGCTTGAACGGGAAAGCCGGCGAAGCTGGCAGAAAGGA 6066  
QY 7736 agggaaagaaacgaagagcgcgctagagcgctgtgcaaaagtgaagcgtaagcttgcg 7795  
| | | | |  
Db 6067 AGGGAAGAAAGCGAAGAGAGCGGCGCTAGGGCGCTGGCAAGTGTAGCCGTACGCTGGC 6126  
QY 7796 cgtacccacacacccgcgcgtctaaatgacgctacagaggcgctgcgcacatctgcg 7855  
| | | | |  
Db 6127 CGTAACCAACCAACACCGCGCGCTTAATGCGCGCTACAGGGCGGCTGCGCATTTGCGC 6186  
QY 7856 attcaagctaacgaactgttgggaagcgatcgatgcggtcttcgctattacgca 7915  
| | | | |  
Db 6187 ATTGAGGCTAGCAACTGTGGGAAGGGCGATGGTGGCGGCTCTTCGCTATTACGCCA 6246  
QY 7916 gctggc 7921  
| | |  
Db 6247 GCGTGC 6252

RESULT 14  
US-08-893-327-17  
Sequence 17, Application US/08893327  
Patent No. 6020192  
GENERAL INFORMATION:  
APPLICANT: Zolotukhin, Sergei  
APPLICANT: Hauswirth, William W.  
APPLICANT: Muzyczka, Nicholas  
TITLE OF INVENTION: Humanized Green Fluorescent Protein  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P. O. Box 4433  
CITY: Houston  
STATE: TX  
COUNTRY: USA  
ZIP: 77210-4433  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/893,327  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/588,201  
FILING DATE: 18-JAN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Kitchenell, Barbara S.  
REGISTRATION NUMBER: 33,928  
REFERENCE/DOCKET NUMBER: DFLA-062\KIT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6280 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 988..1728  
US-08-893-327-17

Query Match 37.6% Score 2985.2; DB 3; Length 6280;  
Best Local Similarity 99.8% Pred. No. 0;  
Matches 3000; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY 4916 agatctaggaaacccttaatgatgatgttggcaactccctctctgcgctcgtctca 4975  
| | | | |  
Db 3277 AGATCTAGGAACCCCTTAGTGAATGAGATTGGCCACTCCCTCTGGGCGCTGCTGCCTCA 3336  
QY 4976 ctgaagcgcccggaagacccggcgctcgaggacctttgtctgcgcgcgtcaatga 5035  
| | | | |  
Db 3337 CTGAGGCGCCCGGGCAAGCCGGGCGCTGGGCGCACTTTGTGTGGCCCGGCTCACTGA 3396  
QY 5036 ggcagcgagcgcgagagagagtgatgcgaacccccccccccccccctgcagccagc 5095  
| | | | |  
Db 3397 GCGACGAGCGCGCAGAGAGGAGTGGCCAAACCCCCCCCCCCCCCTTGACACCC---C 3453  
QY 5096 tgcattaatgaatcgcgcaacgcgcgggagagagcggtttgcatttggcgctcttcg 5155  
| | | | |  
Db 3454 TGCATTAAATGAATCGGCCAACCGCGGGGAGAGGGGTTTCGATTATGGCGCTCTTCG 3513  
QY 5156 ctctcctgcctacgtactcgtctgcgtcgtctgcgttcgctgcgcgagcggatcagctc 5215  
| | | | |  
Db 3514 CTCTCTCGCTCACTGACTGCTGCGCTCGTGGCTGCTGCGGCTGCGGCGATCACTC 3573  
QY 5216 actcaaggcggttaatacgttatccacaagaatacaggagataacgcggaaagacatgt 5275  
| | | | |  
Db 3574 ACTCAAGCGGTAATACGTTATCCACAGATCAAGGAGATACGAGAAAGAACATGT 3633  
QY 5276 gagcaaaagcgcaagaaagccaggaacgtaaaagccgcttgcgcttcttc 5335  
| | | | |  
Db 3634 GAGCAAAAGCGCAGCAAAAGCGCAGAACCGTAAAAAGCGCGCTTCTGCGCTTTTTC 3693  
QY 5336 ataggctccgccccctgaagcaatcaaaaaatcgacgtcgaatcgaagtgagcgaa 5395  
| | | | |  
Db 3694 ATAGGCTCGCGCCCGCTACGAGCATCAAAAATCGACGCTCAAGTCAGAGTGGGGAA 3753  
QY 5396 accgacagagataaagataacagcgatctccctcggaagctccctcgtgcgctc 5455  
| | | | |  
Db 3754 ACCGACAGAGATTAAAGATACAGGCGTTTCCCTGGAAGCTCCCTGCTGGCTCTC 3813  
QY 5456 ctgttcgacacctgcgcttaccgatactcgtcgcgcttctcccttcggaagcgctg 5515  
| | | | |  
Db 3814 CTGTTCCGACCGCTCGCGCTTACCGGATACGTCGCTGCGCTTCTCCTCGGGAAGCGTGG 3873  
QY 5516 cgcttctcaatgctcaagcgtgaagtatctcaagttcggtgaagtgctgcctcaagc 5575  
| | | | |  
Db 3874 CGCTTTCGAAGCTCAGCGCTGTAGTATCTCAGTTCGCTAGCTGCTTCGCTCCAGC 3933  
QY 5576 tgggctgtgtgacgaaccccccgctcagccgcgacgcgtgcgttatccgtaactatc 5635  
| | | | |  
Db 3934 TGGGCTGTGTGACGAACCCCGCTTCA6CCGACCGCTGCGCTTATCCGTAATATC 3993  
QY 5636 gtcttgatgccaaaccggttaagaacagactatcgccactgycagcagccaatgtaaca 5695  
| | | | |  
Db 3994 GCTTGAGTCCAAACCGGTAAGACAGCACTTATGCGCCACTGGACGACCACTGTAAACA 4053  
QY 5696 ggaatgaagcagagatgtagcggtgtctacaagatcttgaagtgtgtgcctaact 5755  
| | | | |  
Db 4054 GGATTAGCAGAGCGAGGTATGAGCGGCTCTTAAGATTTTAAAGTGGTGGCTTAAC 4113  
QY 5756 acggtacacagaaaggaacagatttggatctcgcgtcgtcctaagccagttacactcg 5815  
| | | | |  
Db 4114 ACGGCTACACTAGAGAGAGAGATTTGGTATCTGGCTCTGCTTAAGCCAGTTACCTTCG 4173  
QY 5816 gaaaaagattgtagctcttgatccggaacaaacacacccgctgtagcgtgtgtttt 5875  
| | | | |  
Db 4174 GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACACACCGCTGTGAGGGGTGTTT 4233  
QY 5876 ttgttgcagcagcagatcgcgcgcagaaaaaagatctcaagaagatcccttgatct 5935  
| | | | |  
Db 4234 TTGTTGCAACAGCAGATTACGGCAGAAAAAAGAGTCTCAAGAAAGATCTTTGATCT 4293  
QY 5936 ttctacggggtctgaagctcagtggaacgaaactcaacgtttaaggatcttgtcatga 5995  
| | | | |  
Db 4294 TTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGATTTTGTCTATGA 4353

QY 5996 gattatcaaaaagatcttcacccatagatccctttaaataaaatgaagttttaaaca 6055  
|||||  
Db 4354 GATTATCAAAAAGATCTTCACCTAGATCCCTTTAAATTAATAATGAAGTTTAAATCA 4413  
|||||  
QY 6056 tctaaagtataatgaagtaaaacttgctgaagttacaaatgcttaacgaatgaagcac 6115  
|||||  
Db 4414 TCTAAAGTATATATGAGTAAACTGTGCTGACACTTACCAATGCTTAATCAGTAGGAC 4473  
|||||  
QY 6116 ctactcaagcatctgtctatcttcttcatcatcatgaatgtgcctgactcccgctgtaga 6175  
|||||  
Db 4474 CTATCTACGCACTGTCTATTTGTTATCTCATAGTATGCTGACTCCGCTCTGTAGA 4553  
|||||  
QY 6176 taactacgatatacggagagggcttaccatctggcccaagtgtcgtcaatgatalaaccgagacc 6235  
|||||  
Db 4534 TAATACGATACGGAGAGGGCTTACCATCTGGCCCAAGTGTGCAATGATACGGAGAGCC 4593  
|||||  
QY 6236 caagctcaacggctccagattatcagcaataaaccagccagccggaaggcgagcga 6295  
|||||  
Db 4594 CACGCTACCGGCTCCAGATTATCAGCAATTAACAGCCAGCCGGAAGGGCCGAGCGCA 4653  
|||||  
QY 6296 gaagtgtccctgaacttaccgctccatccagtcctatattgttctgcccggaaagcta 6355  
|||||  
Db 4654 GAATGTGCTCGACACTTTATCGGCTCATCCAGTCTATTATTTGTCGGGAGACTA 4713  
|||||  
QY 6356 gaagtaagtagtcgcagcttaataagtttgcgcaacgttgttgcattgtcacaagcatcg 6415  
|||||  
Db 4714 GAGTAAGTAGTTCGCCAGTTAATAGTTGGCAACGTTGTTGCCATTGCTACAGCATTCG 4773  
|||||  
QY 6416 tgtgtcaagctcgctgttggataggtctcatcagctccggttcccaacgatcaagc 6475  
|||||  
Db 4774 TGGGTGACGCTCGTCTGTTGGTATGGCTTCATTCAGCTCGGTGCCAAGATCAAGGC 4833  
|||||  
QY 6476 gaattacatataccccaatgtatgtgcaaaaagaggttaagctcttctggtccctcagtcg 6535  
|||||  
Db 4834 GAGTTACATGATCCCATGTTGTCAAAAGAGGGTTAGCTCTTGGTCTCGATTCG 4893  
|||||  
QY 6536 ttgcagaagaatagttgcccgaqgttatacactacatggtatgycagacactgacataat 6595  
|||||  
Db 4894 TTGTCAGAGTAAGTTGGCCGCACTGTTATCATCATGTTATGGCAGCACTGCATATTT 4953  
|||||  
QY 6596 ctcttaactgtcatatgcatccgtaagaatgtcttctgtgaactgtgtgaactcaacagt 6655  
|||||  
Db 4954 CTCTTACTGTATCCCATCCGTAAGATGCTTTCTGTGACGTGGAGTACTCAACCAAGT 5013  
|||||  
QY 6656 catctcagaagaatgtatgtatgagcgacagagttgtcttgcggcgctcaataagagata 6715  
|||||  
Db 5014 CATTTCTGAGTAATGATGATGGGCGACCGAGTTGCTTGGCCGGCTCATATAGGGATA 5073  
|||||  
QY 6716 ataccgycacacatagacagaactttaaagttgcatcatatgtaaaaaagttcttcgggac 6775  
|||||  
Db 5074 ATACCGGCGCATAGCAGAACTTTAAAGTGTCTCATCATTTGGAACGTTTCTTGGGGC 5133  
|||||  
QY 6776 gaaacctcagaagatcttaccgctgttgaagatccagttcgatgtaaaccaactcgtgac 6835  
|||||  
Db 5134 GAAAACTCTCAAGATCTTACCGCTGTGAGATCCAGTTCATTAACCCACTGTTGCAC 5193  
|||||  
QY 6836 ccaactgatcttcaagcatcttcttacttccacagcgtttctgggtgagcaaaaaaagaa 6895  
|||||  
Db 5194 CCAACTGATCTTCAGACTCTTTTACTTTACACAGCGCTTTCTGGGTGAGCAAAAACAGAA 5253  
|||||  
QY 6896 ggcacaaatgcgcgcaaaaaggaataagggcgacacggaatgttgatactactct 6955  
|||||  
Db 5254 GGCAAAATGCGCGCAAAAAGGAATAGGGGAGACAGGAATGTTGAATCTCATCTCT 5313  
|||||  
QY 6956 tccctttcaatatattatgaagcatcttacaaggttatgtctcaatgaagcgatacatat 7015  
|||||  
Db 5314 TCCATTCTTCAATATATTATGAAGCAATTATCAGGGTTATTGTCATGAGGGGATACATAT 5373  
|||||  
QY 7016 ttgaatattatagaaaaaataaacaataagagtttcgacgcaaatctcccgaaaaagtc 7075  
|||||  
Db 5374 TTGATATATTATTAGAAAAATTAACAAATAGGGGTTCCGCGACATTTTCCCGAAAAATGTC 5433  
|||||  
QY 7076 caactgagcttaagaacaatatatcatgacattaacctataaataagcgatca 7135  
|||||

Db 5434 CACGTACGCTAGAAACCATTTATTCATGACATTAACTATAAAAATAGGGTATCA 5493  
|||||  
QY 7136 cgagggcccttctgtctcgcgcgcttgcgtgatgaacgttgaacacttgacatcgacgc 7195  
|||||  
Db 5494 CGAGGCCCTTTCGTCGCGCGTTCGGTGTATGACGGTGAACCACTGTGACATGACG 5553  
|||||  
QY 7196 tcccgagaagcggtacagactgtctgttaagcgggtgcgcgggaagaacggccgtgaag 7255  
|||||  
Db 5554 TCCGGAGACGGTACAGCTGTGTATAGCGATGCGGGAGAGCAAGCCGTCAGG 5613  
|||||  
QY 7256 ggcgctcaagcgggtgtctgycgggtgtctgagcgttcaactatgycgacatcaagacaga 7315  
|||||  
Db 5614 GCGGTCAGCGGGGTGTGGGGGTGTGGGGCTGTGCTTAATATGCGGATCAAGCAGA 5673  
|||||  
QY 7316 ttgactgaagatgacacatatgcggtgtgtaataacgcacagatgcgttaagaagaat 7375  
|||||  
Db 5674 TTGTACTGAGAGTGCACCAATATGCGGTGTAATTAATGCGCTTAATTTTGTTA 5733  
|||||  
QY 7376 accgcatcaagaaatgttaaacgttaattgttgaatttcgcgttaaatgttga 7435  
|||||  
Db 5734 ACCGATCAGGAATTTTAACGTTAATATTGTTAATTCGCTTAATTTTGTTA 5793  
|||||  
QY 7436 aatcagctatlttlltaacaaatagccgaatccgcaaatccctataaatcaaaaaga 7495  
|||||  
Db 5794 AATCAGCTCATTTTAAACCAATAGCGCAATGCGCAAAATCCCTTAATTAATCAAAAGA 5853  
|||||  
QY 7496 atagaacgaagataggttgaatgttcttccagtttgaacaagaatgcactataaagaa 7555  
|||||  
Db 5854 ATAGCCGAGATAGGTTGATGTTGTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAA 5913  
|||||  
QY 7556 cgttgactccaagtcgaagggcgaaaaacggttatacagggcgaatgcccactagtg 7615  
|||||  
Db 5914 CGTGACTTCAACGTCGAAGGGCGAANAACCGTATACAGGGCATGCCACTACGTGA 5973  
|||||  
QY 7616 accatcacccatcaagatttltggttcgaggttgcgtaacgtaaacacttaaatcggaaacc 7675  
|||||  
Db 5974 ACCATCACCCCTAATCAAGATTTTGGGGTGCAGGTGCGTAAGACACTTAATGGAACCC 6033  
|||||  
QY 7676 taaaggaagcccccgttlttagacttgaacggggaagcggcggaagcttgcgagaagaa 7735  
|||||  
Db 6034 TAAAGGAGGCCCGGATTTGAGCTTACGCGGGGAAACCGCGCAACTGGCGAANAAGGA 6093  
|||||  
QY 7736 agggaaagaagcgaagaagcggcgcttagggcgcttgcgaagtgtaacggctgaagcgcg 7795  
|||||  
Db 6094 AGGGAATAAAGCGAAGAGAGCGGCGCTTAGGGCCCTGCAAGTGTACGGGTACGCTGCG 6153  
|||||  
QY 7796 cgttaaccacacacccgcgcgcttlaatgycgcgttcaagggcgctgcgcacatlcgcc 7855  
|||||  
Db 6154 CGTAACCAACACACCCCGCGGCTTAATGCGCGCTACAGGGCGCGTCCGCCATTCGCC 6213  
|||||  
QY 7856 attcaggctacgcgaactgttgggaagggcgatcgggtgcggcctctcgtataagcga 7915  
|||||  
Db 6214 ATTCAAGGCTACGCAACTTGGGAAGGGCATCGTGCGGGCCCTTCCTATTACGCCA 6273  
|||||  
QY 7916 gctcgcg 7921  
|||||  
Db 6274 GGCTGC 6279  
|||||

RESULT 15  
US-08-893-327-19  
; Sequence 19, Application US/08893327  
; Patent No. 6020192  
; GENERAL INFORMATION:  
; APPLICANT: Zoletukhin, Sergei  
; APPLICANT: Hauswirth, William W.  
; APPLICANT: Muzyczka, Nicholas  
; TITLE OF INVENTION: Humanized Green Fluorescent Protein  
; TITLE OF INVENTION: Genes and Methods  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee

```

1 STREET: P. O. Box 4433
2 CITY: Houston
3 STATE: TX
4 COUNTRY: USA
5 ZIP: 77210-4433
6 COMPUTER READABLE FORM:
7 MEDIUM TYPE: Floppy disk
8 COMPUTER: IBM PC compatible
9 OPERATING SYSTEM: PC-DOS/MS-DOS
10 SOFTWARE: Patent In Release #1.0, Version #1.30
11 CURRENT APPLICATION DATA:
12 APPLICATION NUMBER: US/08/893,327
13 FILING DATE:
14 CLASSIFICATION: 514
15 PRIOR APPLICATION DATA:
16 APPLICATION NUMBER: US 08/588,201
17 FILING DATE: 18-JAN-1996
18 ATTORNEY/AGENT INFORMATION:
19 NAME: Kitchell, Barbara S.
20 REGISTRATION NUMBER: 33,928
21 REFERENCE/DOCKET NUMBER: UFLA:062\KIT
22 TELECOMMUNICATION INFORMATION:
23 TELEPHONE: (512) 418-3000
24 TELEFAX: (713) 789-2679
25 INFORMATION FOR SEQ ID NO: 19:
26 SEQUENCE CHARACTERISTICS:
27 LENGTH: 6280 base pairs
28 TYPE: nucleic acid
29 STRANDEDNESS: single
30 TOPOLOGY: linear
31 FEATURE:
32 NAME/KEY: CDS
33 LOCATION: 988..1728
34 US-08-893,327-19

```

|                            |        |               |       |                   |
|----------------------------|--------|---------------|-------|-------------------|
| Query Match                | 37.68; | Score 2985.2; | DB 3; | Length 6280;      |
| Best Local Similarity      | 99.88; | Pred. No. 0;  |       |                   |
| Matches 3000; Conservative | 0;     | Mismatches    | 3;    | Indels 3; Gaps 1; |

|    |      |  |       |
|----|------|--|-------|
| OY | 4916 | agatcaggaaacccctcagtatgagcttggccaatccctctctcgcgtccgcgca          | 49175 |
| Dd | 3277 | ACATcAGGAACCCCTtAGtATgAGtTGGGCACTCCCTCTCTGCGGCTCGCTCGCTCA        | 33366 |
| OY | 4976 | ctcgaagccgcccgcgcgaagcccgcggtctcggtcgacetttgtctgcgcgcctcagtga    | 50355 |
| Dd | 3337 | CTtAGAGCCCCCGGCGCAAAACCCGGGCGGTGCGGCGACCTTtTGtCGCGCGGCTCAAGTGA   | 33966 |
| OY | 5036 | gggaagagacgcgcgaagagaggttggccaaccccccccccccccttcgcacccaagc       | 50955 |
| Dd | 3397 | GGGAGGAGACGGCCAGAGAGAGGTGGCCAAACCCCCCCCCCCCCCTtCGACGC---C        | 34533 |
| OY | 5096 | tgcattaatgaatctcgcccaacgcgcggggagagcggttttgcgatttggcgaccttcgcg   | 51555 |
| Dd | 3454 | TGCATTtAATGAATCGGCGCAACGCGGGGGAGAGCGGTTtTGCATTtGGGCGCTTTCG       | 35133 |
| OY | 5156 | ctctctcgctcactgaactgcgtcgtcgtgttcgttcggtcgtgcgcgcgcggtatcgaactc  | 52155 |
| Dd | 3514 | CTTCTCGCTCACTGAATCGCTCGCTCGCTCGGTCTGCTTCGGCTCGCGGCGAGCGGTATCACTC | 35733 |
| OY | 5216 | actcaaaagcggtgaatcaggtttatcccaagaatcaagggatcgcgcgaagaagacatgt    | 52757 |
| Dd | 3574 | ACTCAAAAGCGGTATtACGTTtTTCACAAATtCAGGGGATtAACGCAGGAAGAACAATGT     | 36333 |
| OY | 5276 | ggcgaaaagccgcagcaaaaagccaggaaccgttaaaaaagccgcgttgtcgtgcttttcc    | 53355 |
| Dd | 3634 | GAGCAAAAAGCCAGCAAAAAGCCAGGAACCGTtAAAAAGCCGCGCTGTGCGCTTTTTC       | 36933 |
| OY | 5336 | ataagctccgcccccttgaagagcatcaaaaaaatctgaagctcaagtcagaaggtgcgaa    | 53955 |
| Dd | 3694 | ATtAGGCTCCGCCCCCTCGAGAGCACTtCCAAAATtCACGCTCAAGTtCAGAGGTGGCGAA    | 37533 |
| OY | 5396 | accgcgcagagcattaaaagaatacagagcgttttcccccttgaagctccctcgttgcgttc   | 54555 |

|    |      |   |      |
|----|------|---|------|
| Db | 3754 | ACCAGCACGGACCTATAAAGATACCAGGCGTTTCCCTCCCGAAGCTCCCTCGTGGCCTCTC   | 3813 |
| Qy | 5456 | ctgttcacgaacctgacgcttaaccggaataactgttcgcgccttccctctcggaagcgtg   | 5515 |
| Db | 3814 | CTGTTCGCAACCCCTCGCGCTTACGGATACCTGTCCGCCTTCCCTTCCGGAACGCTGG      | 3873 |
| Qy | 5516 | cgcttctcaatgctcaacgctgtgaagratctoaagttcgtgtgtagtgcgttgcctcaacg  | 5575 |
| Qy | 5576 | tgggcgtgtgcacgaacccccgttcaagccgaacgcgtgcgcctatccgylaacta        | 5635 |
| Db | 3874 | CGCTTTCGAATGCTCAACGGCTGTAAGTATCTCAATTCGTTGGGTGAAGTCGTTCCCTCAAGC | 3933 |
| Qy | 5636 | gtcttgagttccaacccggtlaagacaagaactatcgcgaactgcaagcgaactgyla      | 5695 |
| Db | 3994 | GCTTTGACTCTCAACCCGCTAAGACACACACTTACGGCACACTGGCAGCAGCACTGGTAA    | 4053 |
| Qy | 5696 | ggattagagagcgaggtatgtaagcggtgtcaacagttctctgaagttg               | 5755 |
| Db | 4054 | GGATTAGCAGAGCCAGGATATGTACGGCGGTCTACAGATTTCTTGAAGTGTGGCTAACT     | 4113 |
| Qy | 5756 | acgcgtacacataagaagacagtalttgria                                 | 5815 |
| Db | 4114 | ACGGCTACACTAGAGAGCAAGTATTTGGTATCCGCTCTGTGAAGCCAGTTACTCTTG       | 4173 |
| Qy | 5816 | gaaaaagagttgtagctcttgatctcggaaaaaaacgcgctgtgtagcgtgtttt         | 5875 |
| Db | 4174 | GAAAAAGGTTGGTAGCTCTTGATCTCCGGCAACAACAACACCGTGTAGCGGTGGTTTTT     | 4233 |
| Qy | 5876 | tgtttgcgaagcagcagaattacnagcgcaaaaaaagatcacaagaatctccttga        | 5935 |
| Db | 4234 | TTGTTTGCAGCACGACAGATTACGGCGCAGAAAAAAGATCTCAAGAAATCTTTGATCT      | 4293 |
| Qy | 5936 | tttctaaggggtctcgaagcccgatggaagcaaacccaagttaaggtatltgtgtca       | 5995 |
| Db | 4294 | TTTCTACGGGTCTGACGCTAGTGGAGCAACAACCTACGTTAAGGATTTGGTCA           | 4353 |
| Qy | 5996 | gattacaaaaaagatctcaacctagatccctttaaattaaaaa                     | 6055 |
| Db | 4354 | GATTATCAAAAAGGATCTTCACTCCAGTCTTTTAATTAATAAAGATTTTAATCA          | 4413 |
| Qy | 6056 | tctaaaglatatagtaaactgtgtcgaacgttacaatgcttaatcagttgagcac         | 6115 |
| Db | 4414 | TCTAAAGTATTAAGAGTAAACTTGGTCTACAGTTACCAATGCTTAACTAGTGA           | 4473 |
| Qy | 6116 | ctatctcaagcgaatgtgtctatcttcttcgttcaacagtagtgcgtgaacccccgtgtga   | 6175 |
| Db | 4474 | CTATCTACGCACTGTCTGTATTTGGTTCATCCATAGTGTGCTGACTCCCGCTCGTGA       | 4533 |
| Qy | 6176 | taactacgatacggagaggtctaacatctcgtgcccaagtcgtgaatgataccgcga       | 6235 |
| Db | 4534 | TAACTACATATCGGAGAGGCTTACATCTGGCCCCAGTCTGCAATGATACCGCAGAC        | 4593 |
| Qy | 6236 | caagctaacccggtccgaagtttatcgaataaacaagcgaacggaagggcggaagca       | 6295 |
| Db | 4594 | CACGCTCACCGGCTCCAGATTATACACAAATTAACAAGCAGCGCGGAAGGCGCAGAC       | 4653 |
| Qy | 6296 | gaagtgctcgtcaacttatccgcctcccaatcagctctaatgtgttcggaagata         | 6355 |
| Db | 4654 | GAAAGGTGCTGCAAATTTATCCGGCTCATCCAGTCAATTAATTTGTTCCGGGAAGCTA      | 4713 |
| Qy | 6356 | gagtaagtagtctgcgaagtaatagtttcgcaagctgtgtgccatgtctacagga         | 6415 |
| Db | 4714 | GAGTAAAGTAGTTCCCGCAATTAAATAGTTTGGCAACGTTGTCCATTTGTACAGCATCG     | 4773 |
| Qy | 6416 | tgtgtgaacgctcgtgtgtgtgataagcttcatcgaagctccggttcccaagataaagc     | 6475 |
| Db | 4774 | TGTGTCAAGCTCGTGTGTGTGTAAGGCTTCATTCAAGCTCCGGTTCCACAGATCAAGGC     | 4833 |
| Qy | 6476 | gagtaacatgtctcccaatgtgtgcaaaaaagcttagtctcttgcgtctcccgatcgt      | 6535 |

```

Db 4834 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCCTCGTCTCCGATCG 4893
QY 6536 ttgtcaagaatagtttgccgcaagtggtatcactcaatggtatggaacacatgacataatt 6595
Db 4894 TTGTGAGAAATAGTGTGGCCGAGTGTATCATCTCATGTTATGCGACACTGCATATATT 4953
QY 6596 ccttactgtcatgccaatccgtaagatgctttctgtaactggtgagtaactcaacaagt 6655
Db 4954 CTCTTACTGTCATGCCATCCGTAAGATGCTTTCTGTGACTGGTGAGTACTCAACCAAGT 5013
QY 6656 cacttgaagaatagtgtatgctggcgcaacgaattgctcttgccggcggtcaatacggata 6715
Db 5014 CATTTGAGAAATAGTGTATCCGGCGAGCCGAGTTGCTTCCCGCGTCAATCGGGATA 5073
QY 6716 ataccgpgccacataagagaacttaaaatgctcatcatggaataaactgtctcggggc 6775
Db 5074 ATACCGCGCCACATAGCAGAACTTTAAAGTGCCTCATCATTGGAATAACGTTCTTCGGGGC 5133
QY 6776 gaaaactctcaagatcttaccgcgtgtgagatccagltcagatgtaaccactcgtgcac 6835
Db 5134 GAAACTCTCAAGATCTTACCGCTGTGAGATCCAGTTGAGTGAACCCACTCGTGAC 5193
QY 6836 ccaactgatactcaagcatcttacttcaacagcgltctggtgagcaaaaaacagaa 6895
Db 5194 CCAACTGATCTCAGCATCTTTACTTTCACGACGTTCTCGGGTGAAGCAAAACAGGAA 5253
QY 6896 ggcgaatgcgcaaaaaaggaataaggcgcaacgaaatgtgaataatcatctc 6955
Db 5254 GGCAAATGCGCAAAAGGAATAGGGGACGCAAGAAATGTGAATCTATCTCT 5313
QY 6956 tccctttcaatactatgaagcatctatcaaggttatgtcctcagtcagcgagatacat 7015
Db 5314 TCCTTTTCAATATATATGAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT 5373
QY 7016 ttgaatgattatgaaaaaatacaaaataggggtccgcgcacattccccgaaaatgctc 7075
Db 5374 TTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGGCACATTTCCCGGAAAAAGTGC 5433
QY 7076 caccctgaagcttaagaagaacattatatacgaacttaactataaaatagcgfataca 7135
Db 5434 CACCTTACGCTTAAGAAACCATTTATATCATGACATTAACTTAATAAATAGCGGATCA 5493
QY 7136 cgaagcccttcgtctcgcgcgttctggtgatgacggtgaaaaactctgacacatgcagc 7195
Db 5494 CGAGGCCCTTTGCTCGGGCGTTTGGGTGATGACGTTGAAMCCTGTGACACATGACG 5553
QY 7196 tcccggaagacggtcacagcttgcgtgaagcggaatgcgggagcagaacagccgtcagg 7255
Db 5554 TCCCGGAGACGGTCAACGCTTGTGTAAAGCGGATCCGGGAGCAGACAAAGCCGTCAGG 5613
QY 7256 ggcgctcaagcggtgttgccgggtgtccggggtcgttaactatgcgcatcaagaacaga 7315
Db 5614 GCGCGTACGGGGGTGTGGGGGTGTGGGGGTGAGCTTAACCTATGCGCATCAGACAGAA 5673
QY 7316 ttgtaactgagatgacacatagcgtgtgaaataccgcaacagatcgtgaagagaanaat 7375
Db 5674 TTGTACTGAGAGTGCACCATATGCGGTGTGAATACCGCACAGATGCTTAAGAGAAAT 5733
QY 7376 accgcatcaggaatgttaaacgtttaatcttgttaaaatlcgcgttaaatlttgtta 7435
Db 5734 ACCGCAATCAGGAATGTAAACGTTATATTTGTTAAATTCGCGTTAAATTTTGTGA 5793
QY 7436 aatcagctcatctttaaacaatagcgcaaaatcgcaaaatccctataaatcaaaaaga 7495
Db 5794 AATCAGCTCATTTTAAACCAATAGCCGGAATCGCAAAATCCCTTATTAATCAAAAGA 5853
QY 7496 atagaccgagataggttgagtggttctcaggttgaacaagaatccactataaagaa 7555
Db 5854 ATAGACCGAGATAGGTTAGTGTCTTCCAGTTTGAAACAAAGTCCACTATTAAAGAA 5913
QY 7556 cgtgactccaacgltcaaaaggcgaaaaacgltctatcaggcgatgcccactacgtga 7615
Db 5914 CGTGGACTCCAAAGTCAAGGGGGAACCGCTATCAGGGCGATGGCCCACTACGTGA 5973

```

```

QY 7616 accatcacccatacaagtttttgggtgcaggtgcccgtlaaagcaactaaatcgaaccc 7675
Db 5974 ACCATCACCTTAATCAAGTTTGGGGTCAAGTCCGTAAGACCTAATATGGAACCC 6033
QY 7676 taaaggagccccgatttagaacttgaacggtggaagacggtgcaagcgtgcaagaaagga 7735
Db 6034 TAAAGGAGCCCCGATTTAGACTTGAACGGGGAAGCCGGCAACGTGGCGAAGAAAGGA 6093
QY 7736 agggaaagaaagcgaagagcggtcgttagggcgtgcaagtgtgtagcggtcaacgtcg 7795
Db 6094 AGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGAGCAAGTGTAGCGGTACGCTGGC 6153
QY 7796 cgttaaccacacaccccgccgcttaatgagccgctacagaaggcggtcgcccatcggcc 7855
Db 6154 CATTACCAACACACCCCGCGCTTATATGCGCGCTTACAGGGCGCTCGGCCATTTCGCC 6213
QY 7856 attcaggctaacgaactgttggaagggcgatcgtgtcggtcctctcgtatlaagcga 7915
Db 6214 ATTCAAGGCTACGCAACTGTTGGGAAGGGGATGCGTGGGGCTCTTGGCTATTACGCCA 6273
QY 7916 gctggc 7921
Db 6274 GGCTGC 6279

```

Search completed: November 18, 2001, 04:50:04  
Job time: 35006 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 17, 2001, 19:07:48 ; Search time 919.72 Seconds

(without alignments)  
5423.445 Million cell updates/sec

Title: US-09-689-430-1

Perfect score: 7944  
Sequence: 1 tggcactccctctctgcgc.....aggggggggggggggggggt 7944

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: N.Geneseq.0601:\*  
2: /SIDSR/gcgdata/geneseq/geneseqn/NA1980.DAT:\*  
3: /SIDSR/gcgdata/geneseq/geneseqn/NA1981.DAT:\*  
4: /SIDSR/gcgdata/geneseq/geneseqn/NA1982.DAT:\*  
5: /SIDSR/gcgdata/geneseq/geneseqn/NA1983.DAT:\*  
6: /SIDSR/gcgdata/geneseq/geneseqn/NA1984.DAT:\*  
7: /SIDSR/gcgdata/geneseq/geneseqn/NA1985.DAT:\*  
8: /SIDSR/gcgdata/geneseq/geneseqn/NA1986.DAT:\*  
9: /SIDSR/gcgdata/geneseq/geneseqn/NA1987.DAT:\*  
10: /SIDSR/gcgdata/geneseq/geneseqn/NA1988.DAT:\*  
11: /SIDSR/gcgdata/geneseq/geneseqn/NA1989.DAT:\*  
12: /SIDSR/gcgdata/geneseq/geneseqn/NA1990.DAT:\*  
13: /SIDSR/gcgdata/geneseq/geneseqn/NA1991.DAT:\*  
14: /SIDSR/gcgdata/geneseq/geneseqn/NA1992.DAT:\*  
15: /SIDSR/gcgdata/geneseq/geneseqn/NA1993.DAT:\*  
16: /SIDSR/gcgdata/geneseq/geneseqn/NA1994.DAT:\*  
17: /SIDSR/gcgdata/geneseq/geneseqn/NA1995.DAT:\*  
18: /SIDSR/gcgdata/geneseq/geneseqn/NA1996.DAT:\*  
19: /SIDSR/gcgdata/geneseq/geneseqn/NA1997.DAT:\*  
20: /SIDSR/gcgdata/geneseq/geneseqn/NA1998.DAT:\*  
21: /SIDSR/gcgdata/geneseq/geneseqn/NA1999.DAT:\*  
22: /SIDSR/gcgdata/geneseq/geneseqn/NA2000.DAT:\*  
23: /SIDSR/gcgdata/geneseq/geneseqn/NA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | DB | ID        | Description         |
|------------|--------|-------------|--------|----|-----------|---------------------|
| 1          | 6608.4 | 83.2        | 11933  | 21 | AAD00121  | Recombinant adeno   |
| 2          | 5553   | 69.9        | 9164   | 20 | AAH82259  | Beta-domain delete  |
| 3          | 4456.2 | 56.1        | 4999   | 21 | AAD00122  | Recombinant adeno   |
| 4          | 4413   | 55.6        | 4629   | 20 | AAH88293  | Human Factor VIII   |
| 5          | 4411.6 | 55.5        | 9354   | 18 | AAAT73164 | CDNA encoding huma  |
| 6          | 4355.4 | 54.8        | 4629   | 16 | AAO76016  | B-domain deleted F  |
| 7          | 4351.2 | 54.8        | 4670   | 19 | AAV23339  | Human Factor-VIII   |
| 8          | 4321.6 | 54.4        | 4832   | 19 | AAV19581  | Human Factor-VIII   |
| 9          | 4321.6 | 54.4        | 4832   | 19 | AAV13538  | Human Factor VIII   |
| 10         | 4319   | 54.4        | 12445  | 21 | AAA49232  | Vector HSGRenNeo fo |
| 11         | 4173   | 52.5        | 4275   | 10 | AAH90654  | DNA encoding 740 A  |

|    |        |      |       |    |           |                    |
|----|--------|------|-------|----|-----------|--------------------|
| 12 | 4171.6 | 52.5 | 4275  | 9  | AAH80446  | Modified factor VI |
| 13 | 4170   | 52.5 | 4272  | 9  | AAH80447  | Modified factor VI |
| 14 | 4161   | 52.4 | 4373  | 20 | AAH82258  | Beta-domain delete |
| 15 | 4152.2 | 52.3 | 4830  | 9  | AAH81544  | Human Factor VIII- |
| 16 | 4108.8 | 51.7 | 4545  | 9  | AAH80444  | Modified factor VI |
| 17 | 3748   | 47.2 | 5035  | 18 | AAAT69811 | Factor VIII-GB695- |
| 18 | 3708.6 | 46.7 | 4616  | 9  | AAH81545  | Human Factor VIII- |
| 19 | 3678   | 46.3 | 5094  | 21 | AAA49231  | Human Factor VIII- |
| 20 | 3332   | 41.7 | 11846 | 20 | AAH82261  | DNA construct HSG/ |
| 21 | 3311   | 41.7 | 12022 | 20 | AAH82260  | Factor VIII protei |
| 22 | 3189.8 | 40.2 | 4334  | 19 | AAV12142  | Factor VIII protei |
| 23 | 3189.8 | 40.2 | 4334  | 20 | AAH91195  | Homo sapiens facto |
| 24 | 3189.8 | 40.2 | 4334  | 22 | AAH90541  | Porcine factor VII |
| 25 | 2999.4 | 37.8 | 6981  | 21 | AAZ45931  | CDNA encoding porc |
| 26 | 2985.2 | 37.6 | 6253  | 20 | AAH08454  | Nucleotide sequenc |
| 27 | 2985.2 | 37.6 | 6280  | 20 | AAH08455  | AAV vector sequenc |
| 28 | 2985.2 | 37.6 | 6280  | 20 | AAH08456  | AAV vector sequenc |
| 29 | 2985.2 | 37.6 | 6565  | 21 | AAZ45925  | Nucleotide sequenc |
| 30 | 2985.2 | 37.6 | 7054  | 21 | AAZ45927  | Nucleotide sequenc |
| 31 | 2985.2 | 37.6 | 7405  | 21 | AAZ45926  | Nucleotide sequenc |
| 32 | 2947   | 37.1 | 7492  | 21 | AAZ45929  | Nucleotide sequenc |
| 33 | 2836.4 | 35.7 | 5932  | 21 | AAZ45928  | Nucleotide sequenc |
| 34 | 2836.4 | 35.7 | 6142  | 21 | AAZ45932  | Nucleotide sequenc |
| 35 | 2836.4 | 35.7 | 6142  | 21 | AAZ45933  | Nucleotide sequenc |
| 36 | 2836.4 | 35.7 | 6714  | 21 | AAZ45930  | Nucleotide sequenc |
| 37 | 2836.4 | 35.7 | 6924  | 21 | AAZ45934  | Nucleotide sequenc |
| 38 | 2836.4 | 35.7 | 6924  | 21 | AAZ45935  | Nucleotide sequenc |
| 39 | 2830.6 | 35.6 | 10930 | 16 | AAO81226  | Plasmid pm16-1. C  |
| 40 | 2830.6 | 35.6 | 10950 | 16 | AAO81225  | Plasmid pm16. Chi  |
| 41 | 2794.2 | 35.2 | 4118  | 18 | AAAT69188 | Construct pGEM-hTR |
| 42 | 2794.2 | 35.2 | 4118  | 18 | AAAT69189 | Construct pGEM-hTR |
| 43 | 2794.2 | 35.2 | 7175  | 21 | AAA26728  | Plasmid pCALIPNIST |
| 44 | 2794.2 | 35.2 | 7177  | 21 | AAA52989  | Plasmid pCALIPNIST |
| 45 | 2762.6 | 34.8 | 6094  | 18 | AAV05703  | Plasmid pcre-LBDlg |

ALIGNMENTS

|                                      |   |
|--------------------------------------|---|
| RESULT 1                             |   |
| AAD00121                             |   |
| ID AAD00121 standard; DNA: 11933 BP. |   |
| AC AAD00121;                         |   |
| DT 31-JUL-2000 (first entry)         |   |
| DE                                   | Recombinant adeno associated vector construct, PAAV-F8-1.                 |
| XX                                   |   |
| KW                                   | Recombinant Adeno Associated Vector; PAAV-F8-1; human Factor VIII;        |
| KW                                   | hFVIII; hNF-3 albumin promoter; human elongation factor-1alpha; EF1alpha; |
| KW                                   | human growth hormone; hGH; Inverted terminal repeat; ITR; haemophilia;    |
| KW                                   | gene therapy; ds.   |
| OS                                   | Adeno associated virus.   |
| XX                                   |   |
| PN                                   | WO200023116-A1.   |
| PD                                   |   |
| XX                                   |   |
| PD                                   | 27-APR-2000.  |
| XX                                   |   |
| PF                                   | 19-OCT-1999; 99WO-US24495.  |
| XX                                   |   |
| PR                                   | 20-OCT-1998; 98US-0104994.  |
| PR                                   | 24-MAR-1998; 99US-0125874.  |
| PR                                   | 30-JUL-1999; 99US-0364862.  |
| PA                                   | (AVIG-) AVIGEN INC.   |
| XX                                   |   |
| PI                                   | Couto LB, Colosi PC;  |
| XX                                   |   |
| DR                                   | WPI; 2000-339536/29.  |
| XX                                   |   |
| PT                                   | New recombinant adenovirus-associated vector, useful for gene therapy     |



Db 1684 gccccgaatgcagaagtataaaagtcgaatatttgacaatgcccctcagcggattggt 1743  
QY 1737 aggaagtacaaaaagcccgatttatgtgcatacacagatgaacaccttaagacctgtaa 1796  
1744 aggaagtacaaaaagcccgatttatgtgcatacacagatgaacaccttaagacctgtaa 1803  
QY 1797 gctatcagcatgaatcagaagaatcttgggaccttaactttatvgggaagtvggacaca 1856  
Db 1804 gctatcagcatgaatcagaagaatcttgggaccttaactttatvgggaagtvggacaca 1863  
QY 1857 ctgttgattatatttaagaatccaagcagacacataacatctcccccaagatc 1916  
Db 1864 ccgttgattatatttaagaatccaagcagacacataacatctcccccaagatc 1923  
QY 1917 actgattccgcttccttgatttcagaagagattaccaaaagtgtaaacatttgaagat 1976  
Db 1924 actgattccgcttccttgatttcagaagagattaccaaaagtgtaaacatttgaagat 1983  
QY 1977 ttcccaattctgcagagagaataattcaaatataatvgcaagtgcgtgtagaagatg 2036  
Db 1984 ttcccaattctgcagagagaataattcaaatataatvgcaagtgcgtgtagaagatg 2043  
QY 2037 ccaactaatatcagatcctctggtgcgtgcacccgctatacctctagtttcgttaataatg 2086  
Db 2044 ccaactaatatcagatcctctggtgcgtgcacccgctatacctctagtttcgttaataatg 2103  
QY 2097 agaatatagcttcagaagcaatctgcccctctctcatctgcatacaagaatctgtatg 2156  
Db 2104 agaatatagcttcagaagcaatctgcccctctctcatctgcatacaagaatctgtatg 2163  
QY 2157 caaagagaaaccagataatgctcagacaagaagatgcatcctgtttcttgataattgat 2216  
Db 2164 caaagagaaaccagataatgctcagacaagaagatgcatcctgtttcttgataattgat 2223  
QY 2217 gagaaaccgaagctggtgtcccccacagaataataacaagctttcccccacacgcgtgga 2276  
Db 2224 gagaaaccgaagctggtgtcccccacagaataataacaagctttcccccacacgcgtgga 2283  
QY 2277 gtgagcttgaggaatccagagatcccaagcctccacatactgcacagatcaatggtctat 2336  
Db 2284 gtgagcttgaggaatccagagatcccaagcctccacatactgcacagatcaatggtctat 2343  
QY 2337 gttttgatagtttgcaagtgtcgaagttgttgcaagaagtgagatactggtatattcta 2396  
Db 2344 gttttgatagtttgcaagtgtcgaagttgttgcaagaagtgagatactggtatattcta 2403  
QY 2397 agcatlgtgagcaacagactgactctctctgtcttctctctctgataataccttcaaacac 2456  
Db 2404 agcatlgtgagcaacagactgactctctctgtcttctctctctgataataccttcaaacac 2463  
QY 2457 aaaaatgtctatagaagacacactcaacctatcccatctccagagagaactgtctcatg 2516  
Db 2464 aaaaatgtctatagaagacacactcaacctatcccatctccagagagaactgtctcatg 2523  
QY 2517 tcgaatggaaccccaaggtctatagatctcgggtgycacaacactcagaacttcggagaca 2576  
Db 2524 tcgaatggaaccccaaggtctatagatctcgggtgycacaacactcagaacttcggagaca 2583  
QY 2577 ggcataagccgcttactgaagagtttctaagtgtgagacaagaacactggtgattataagag 2636  
Db 2584 ggcataagccgcttactgaagagtttctaagtgtgagacaagaacactggtgattataagag 2643  
QY 2637 gacagttatgaagatatcttcaagcatactgtgataaacaatgagcatctgaacaaga 2696  
Db 2644 gacagttatgaagatatcttcaagcatactgtgataaacaatgagcatctgaacaaga 2703  
QY 2697 agcttctccagaatttcaagaacactagcaactagccaagaacttaatgcccacca 2756  
Db 2704 agcttctccagaatttcaagaacactagcaactagccaagaacttaatgcccacca 2709  
QY 2757 ccagttctgaaagccatacaggaataataactcgtactactctctcagtcagatcaagag 2816  
Db 2710 -----gaaataactcgtactactctctcagtcagatcaagag 2745

QY 2817 gaaattgactatgataatcacatatacagttgaaatgagaagagaatttgacatttat 2876  
Db 2746 gaaattgactatgataatcacatatacagttgaaatgagaagagaatttgacatttat 2805  
QY 2877 gatgaggaatgaaataatcagaagcccccagacttccaagaagaacagacataatttatc 2936  
Db 2806 gatgaggaatgaaataatcagaagcccccagacttccaagaagaacagacataatttatc 2865  
QY 2937 gctcagatgagaagagctctggatataatgagatgagctccccaatgttctgaagaac 2996  
Db 2866 gctcagatgagaagagctctggatataatgagatgagctccccaatgttctgaagaac 2925  
QY 2997 agggctcagaagtgcagtgctccctcagttcaagaagaattgttttccaggaatttaccgat 3056  
Db 2926 agggctcagaagtgcagtgctccctcagttcaagaagaattgttttccaggaatttaccgat 2985  
QY 3057 ggtccttactcagaccctataacggtgagaactaaatgaaacttvggactcctgggg 3116  
Db 2986 ggtccttactcagaccctataacggtgagaactaaatgaaacttvggactcctgggg 3045  
QY 3117 ccatataagagagagaagtttgaagaataatacatatgttaactttccaagaatacagcctct 3176  
Db 3046 ccatataagagagagaagtttgaagaataatacatatgttaactttccaagaatacagcctct 3105  
QY 3177 cgtccctattcctctatacttcaagccttatcttcatagagagaatcagaagcagaagca 3236  
Db 3106 cgtccctattcctctatacttcaagccttatcttcatagagagaatcagaagcagaagca 3165  
QY 3337 gaacctagaanaaaacttgttccaagcctaatagaaacccaacttacttvgaaagtgc 3296  
Db 3166 gaacctagaanaaaacttgttccaagcctaatagaaacccaacttacttvgaaagtgc 3225  
QY 3397 catcatatgagcaccactaaagatgagatttgacagcagaagcctggtttttcttcgat 3356  
Db 3226 catcatatgagcaccactaaagatgagatttgacagcagaagcctggtttttcttcgat 3285  
QY 3357 gttagacctggaanaaaagtgtgcactcagcctggaattgagccctctcgtctgcacact 3345  
Db 3286 gttagacctggaanaaaagtgtgcactcagcctggaattgagccctctcgtctgcacact 3345  
QY 3417 aacacactgaaacctgctcatatgggagacaagtgcagatcagaagaattgctctgttttc 3476  
Db 3346 aacacactgaaacctgctcatatgggagacaagtgcagatcagaagaattgctctgttttc 3405  
QY 3477 accactcttgatgagaccaaagcgtgttacttaacttgaanaataagaaactgcagg 3536  
Db 3406 accactcttgatgagaccaaagcgtgttacttaacttgaanaataagaaactgcagg 3465  
QY 3537 gctccctgcataatccagatcgaagaatcccaactttaaagaagaattatcgccttcacgca 3596  
Db 3466 gctccctgcataatccagatcgaagaatcccaactttaaagaagaattatcgccttcacgca 3525  
QY 3597 atcaatggtcacaataatgatactacactcgtgcttagtaatggtcaggaatcaaaagatc 3656  
Db 3526 atcaatggtcacaataatgatactacactcgtgcttagtaatggtcaggaatcaaaagatc 3585  
QY 3657 cgaatggtatctgtccagcagggcagcaatgaaacttcatcttatcttcaagtggga 3716  
Db 3586 cgaatggtatctgtccagcagggcagcaatgaaacttcatcttatcttcaagtggga 3645  
QY 3717 catgtgtcactgtatacgaanaaaagagagatataaataatgacatgataatcctaccca 3776  
Db 3646 catgtgtcactgtatacgaanaaaagagagatataaataatgacatgataatcctaccca 3705  
QY 3777 ggtgtttttgagaacagttggaatgttacaatccaagctggaatttggcgggtggaatgc 3836  
Db 3706 ggtgtttttgagaacagttggaatgttacaatccaagctggaatttggcgggtggaatgc 3765  
QY 3837 cttaatggcagacatcaactgtctggaatgagacacttttctgtgtacgaagaatg 3896  
Db 3766 cttaatggcagacatcaactgtctggaatgagacacacttttctgtgtacgaagaatg 3825

|    |      |   |      |
|----|------|---|------|
| QY | 3897 | tgctcagactcccttcgggaatgagcttcctcgacaactttagagatttccagattacagcttca | 3950 |
| Db | 3826 | tgctcagactcccttcgggaatgagcttcctcgacaactttagagatttccagattacagcttca | 3885 |
| QY | 3957 | ggacaatatggaacgttggccccaaagctggccagacttcatctccgagatcaaat          | 4016 |
| Db | 3886 | ggacaatatggaacgttggccccaaagctggccagacttcatctccgagatcaaat          | 3945 |
| QY | 4017 | gctctggagacaacaaggacccttctcttgatcaaggtggaatctcgttgcaccaatgatt     | 4076 |
| Db | 3946 | gctctggagacaacaaggacccttctcttgatcaaggtggaatctcgttgcaccaatgatt     | 4005 |
| QY | 4077 | attcagcgcatcaagaccaccgggtgccccgtcagaagattctccagcctctaatctctcag    | 4136 |
| Db | 4006 | attcagcgcatcaagaccaccgggtgccccgtcagaagattctccagcctctaatctctcag    | 4065 |
| QY | 4137 | tttatcatctgatatgattcttgatgggaagaagtgcgagactttatcgagaaattccact     | 4196 |
| Db | 4066 | tttatcatctgatatgattcttgatgggaagaagtgcgagactttatcgagaaattccact     | 4125 |
| QY | 4197 | ggacaacttaatgttctctcttggcaagtttgatcatctgggaataaacaacaatatattt     | 4256 |
| Db | 4126 | ggacaacttaatgttctctcttggcaagtttgatcatctgggaataaacaacaatatattt     | 4185 |
| QY | 4257 | aaccttccaattatgtctgatatacccggtttgcaccaactaatatagcatctgcagc        | 4316 |
| Db | 4186 | aaccttccaattatgtctgatatacccggtttgcaccaactaatatagcatctgcagc        | 4245 |
| QY | 4317 | acccttcgcataatgtgatgtgggtgcttgatattaaagtgtgcagaatgcattgggaatg     | 4376 |
| Db | 4246 | acccttcgcataatgtgatgtgggtgcttgatattaaagtgtgcagaatgcattgggaatg     | 4305 |
| QY | 4377 | gagagtaaaagcaaatatcagatgacagattactgtctcatccactctatcaaatatttt      | 4436 |
| Db | 4306 | gagagtaaaagcaaatatcagatgacagattactgtctcatccactctatcaaatatttt      | 4365 |
| QY | 4437 | ggcaactgtgtctctcttcaaaaagctcgactccaectccaaggagagtaatgctctgaga     | 4486 |
| Db | 4366 | ggcaactgtgtctctcttcaaaaagctcgactccaectccaaggagagtaatgctctgaga     | 4425 |
| QY | 4497 | ccttcaggttgataatccaagaagatgtgctgcaatltgagcttccagagaacataaagaatc   | 4556 |
| Db | 4426 | ccttcaggttgataatccaagaagatgtgctgcaatltgagcttccagagaacataaagaatc   | 4485 |
| QY | 4557 | acaaggagtaactactcctcagggagtaaaatctctgtttacaagaagatltgtaaagagttc   | 4616 |
| Db | 4486 | acaaggagtaactactcctcagggagtaaaatctctgtttacaagaagatltgtaaagagttc   | 4545 |
| QY | 4617 | ctcatctccagagctcaagaatggccatcagtggaactcttcttttcagaatgtgcaagta     | 4676 |
| Db | 4546 | ctcatctccagagctcaagaatggccatcagtggaactcttcttttcagaatgtgcaagta     | 4605 |
| QY | 4677 | aaggtttttttagaggaaatcaagactccttcaacactgttgtagaactctctagaccacag    | 4736 |
| Db | 4606 | aaggtttttttagaggaaatcaagactccttcaacactgttgtagaactctctagaccacag    | 4665 |
| QY | 4737 | ttctcagctgcacactctcgaattcacccccagaagtttgggtgacacgaattgtccctag     | 4796 |
| Db | 4666 | ttctcagctgcacactctcgaattcacccccagaagtttgggtgacacgaattgtccctag     | 4725 |
| QY | 4797 | atggaagttcttggctgctgcagacagaagacctactgactcgacgagttctcttgag        | 4856 |
| Db | 4726 | atggaagttcttggctgctgcagacagaagacctactgactcgacgagttctcttgag        | 4770 |
| QY | 4857 | ggatcggcgaataaagaacagataaagaacgacgggtggttggttcgcttcttgatcca       | 4916 |
| Db | 4771 | -----aataaagaatcagaagctctcagaagatctgtgtgttggcttcttcttgctgcg       | 4819 |
| QY | 4917 | gactatggaacccctagatgattgagtttggcaactcctctcttggcgctcgtctggcaac     | 4976 |
| Db | 4820 | gcccgcaggaacccctagatgattgagtttggcaactcctctcttggcgctcgtctggcaac    | 4879 |
| QY | 4977 | tgaagccgcgccggcgaagaacccggcgcttgcggcacttttgctgcgccggcctcagtgag    | 5036 |

[illegible]

Db 5753 cttaagatataatagtaaaacttgctgacagttacaatgctaaacagtgagccacc 5812  
QY 6117 tatctagcgatcgtctattcttgctatccatagtgccctgactcccgctcgtgtgat 6176  
Db 5813 tatccagcgatccgtctactctcgtctacccaatagtgccctgactcccgctcgtgtgat 5872  
QY 6177 aactacgatacggaggggcttaccatctggcccgagtgctgcgaatgataccgcagacc 6236  
Db 5873 aactacgatacggaggggcttaccatctggcccgagtgctgcgaatgataccgcagacc 5932  
QY 6237 acgctcacccggtctccagattatcatgaacaacacagccagccggaaggccgagccag 6296  
Db 5933 acgctcacccggtctccagattatcatgaacaacacagccagccggaaggccgagccag 5992  
QY 6297 aaggggcccgcgaacttaccctccatccagttctatattatgtgtcccggaagctag 6356  
Db 5993 aaggggcccgcgaacttaccctccatccagttctatattatgtgtcccggaagctag 6052  
QY 6357 agtaagtagtccgcaatgattgtgcgaacggtgtgtccatctgtacagacgtcgt 6416  
Db 6053 agtaagtagtccgcaatgattgtgcgaacggtgtgtccatctgtacagacgtcgt 6112  
QY 6417 ggtgtacgctcgtcgtgtgtgtatgctcattcagctccggtcccaacgataaggcg 6476  
Db 6113 ggtgtacgctcgtcgtgtgtgtatgctcattcagctccggtcccaacgataaggcg 6172  
QY 6477 agttacaatgtaaccccaatgtgtgcgaacgaagggttagtccctcgtcctccgacgt 6536  
Db 6173 agttacaatgtaaccccaatgtgtgcgaacgaagggttagtccctcgtcctccgacgt 6232  
QY 6537 tgtcaagaagaattgtgcccgaatgttatcatcatatgattgtagcgaacacgtacataatc 6596  
Db 6233 tgtcaagaagaattgtgcccgaatgttatcatcatatgattgtagcgaacacgtacataatc 6292  
QY 6597 tcttactgctatgcatcccgtaagaatgcttctgtgtgacggtgtgagtactcaacacgaatc 6656  
Db 6293 tcttactgctatgcatcccgtaagaatgcttctgtgtgacggtgtgagtactcaacacgaatc 6352  
QY 6657 attctggaatgtagttagtgcgcgacccgaattgctcttgcccgccggtcgaataagggttaa 6716  
Db 6353 attctggaatgtagttagtgcgcgacccgaattgctcttgcccgccggtcgaataagggttaa 6412  
QY 6717 taccgcgcacatagcgaacatttaaaagtctcatcatctgtgaaacagttcttcggggcg 6776  
Db 6413 taccgcgcacatagcgaacatttaaaagtctcatcatctgtgaaacagttcttcggggcg 6472  
QY 6777 aaaaactcgaagatcttaccgctgtgtgagatccagttcgaatgtaaccacactcgtgaccc 6836  
Db 6473 aaaaactcgaagatcttaccgctgtgtgagatccagttcgaatgtaaccacactcgtgaccc 6532  
QY 6837 caactgactctcagcgtcttcttacttcacacaggttctcgggttgagcaaaaacaggaag 6896  
Db 6533 caactgactctcagcgtcttcttacttcacacaggttctcgggttgagcaaaaacaggaag 6592  
QY 6897 gcaaaatgcgcgaaaaagggaaataaggcgcaacgcgaaatgtgtgaaatactactcctt 6956  
Db 6593 gcaaaatgcgcgaaaaagggaaataaggcgcaacgcgaaatgtgtgaaatactactcctt 6652  
QY 6957 ccttttcaatatattgaagcatctacaggttatgtctcatagagcgatacatattc 7016  
Db 6653 ccttttcaatatattgaagcatctacaggttatgtctcatagagcgatacatattc 6712  
QY 7017 tgaatgatttagaaaaataaacaataaggggttcgcgcgacattcccccgaaaaagggcc 7076  
Db 6713 tgaatgatttagaaaaataaacaataaggggttcgcgcgacattcccccgaaaaagggcc 6772  
QY 7077 acctgacgtctaaagaaacattatcatagatataaccataaataatagcgatacac 7136  
Db 6773 acctgacgtctaaagaaacattatcatagatataaccataaataatagcgatacac 6832  
QY 7137 gaagcccttcgctcgcgcgcttgcgtgtgatacggtgaaacccctgacacatgcagct 7196  
Db 6833 gaagcccttcgctcgcgcgcttgcgtgtgatacggtgaaacccctgacacatgcagct 6892

QY 7197 cccggagacggtcacacagcttctctgttaagcggatgcccggagacgaacaaagcccgtaagg 7256  
Db 6893 cccggagacggtcacacagcttctctgttaagcggatgcccggagacgaacaaagcccgtaagg 6952  
QY 7257 cgcgtcagcgggtgtgtgcgggtgtccgggttcggttaactatgcgcatcagaacagat 7316  
Db 6953 cgcgtcagcgggtgtgtgcgggtgtccgggttcggttaactatgcgcatcagaacagat 7012  
QY 7317 tgtactgaagatgcacacatctgcgtgtgaaataccgcgacagatgctgtaaggaaata 7376  
Db 7013 tgtactgaagatgcacacatctgcgtgtgaaataccgcgacagatgctgtaaggaaata 7032  
QY 7377 ccgcataaggaattgtgaacggttaatttgtttaaattcgcgttaaatctttgttaa 7436  
Db 7033 -----aaatgtgaacggttaatttgtttaaattcgcgttaaatctttgttaa 7082  
QY 7437 atcagctcatctttaaaccataagccgaaatcgcgaatcccttaataatcaaaagaa 7496  
Db 7083 atcagctcatctttaaaccataagccgaaatcgcgaatcccttaataatcaaaagaa 7142  
QY 7497 tagaacgagataagggttagtgtgttccagtttggaacaagatccactataaagaac 7556  
Db 7143 tagaacgagataagggttagtgtgttccagtttggaacaagatccactataaagaac 7202  
QY 7557 gtggactccacagtcgaaggcggaacacgtctatcaggcgatgccccactaagttaa 7616  
Db 7203 gtggactccacagtcgaaggcggaacacgtctatcaggcgatgccccactaagttaa 7262  
QY 7617 ccatcaccccttaacaaagtcttctgggggtcgaaggctgccgtlaaagcactaaatcggaaacct 7676  
Db 7263 ccatcaccccttaacaaagtcttctgggggtcgaagggtccgtlaaagcactaaatcggaaacct 7222  
QY 7677 aaagggaagcccccgaatttagagcttgaacggggaagccggcggaacgttgcgagaagaagaa 7736  
Db 7323 aaagggaagcccccgaatttagagcttgaacggggaagccggcggaacgttgcgagaagaagaa 7382  
QY 7737 ggggaagaagcgaagaagggcggtcgttagggcgcttggaagcttgaacggcgacgctgcgc 7796  
Db 7383 ggggaagaagcgaagaagggcggtcgttagggcgcttggaagcttgaacggcgacgctgcgc 7442  
QY 7797 gtaaccacacaccccgccgcttaatgctgcgcgtcacaagggcgct 7842  
Db 7443 gtaaccacacaccccgccgcttaatgctgcgcgtcacaagggcgct 7488

RESULT 2  
AAx82259  
ID AAx82259 standard; cDNA; 9164 BP.  
XX  
AC AAx82259;  
XX  
DT 18-AUG-1999 (first entry)  
XX  
DE Beta-domain deleted Factor VIII protein encoding gene (uncorrected).  
XX  
KW Factor VIII protein; gene modification; gene therapy; clinical disorder;  
KW Splicing pattern; RNA processing; gene regulation; beta-domain; human;  
KW ss.  
XX  
OS Homo sapiens.  
XX  
PN W09929848-A1.  
PD 17-JUN-1999.  
XX  
PF 25-NOV-1998; 98WO-US25354.  
XX  
PR 16-JAN-1998; 98US-0071596.  
PR 05-DEC-1997; 97US-0067614.  
XX  
PA (IMMU-) IMMUNE RESPONSE CORP.  
XX

PI Bidlingmaier S, Gonzales JEN, ILL CR, Yang CQ;  
XX  
XX WPI: 1999-385602/32.  
DR P-PSDB; AAY21675.  
XX  
PT Genes and vectors exhibiting increased expression and novel splicing  
PT patterns, useful for expression of, e.g. beta-domain deleted factor  
VIII  
XX  
XX  
PS Example 2; Page 79-89; 123pp; English.  
XX  
XX The invention describes novel genes and vectors exhibiting increased  
CC expression and novel splicing patterns. It provides a gene encoding a  
CC factor VIII protein, that comprises one or more consensus or near  
CC consensus splice sites which have been corrected to increase expression.  
CC The method, DNA sequences and expression vectors can be used to increase  
CC the expression of a gene, especially a Factor VIII gene. Genes containing  
CC modified 5' and/or 3' untranslated regions have optimized expression  
CC levels and tissue-specific expression. The methods are used for  
CC identification and correction of consensus splice sites, addition of  
CC introns, optimization of 5' and 3' untranslated regions and increase in  
CC cytoplasmic RNA accumulation. Hence the DNAs are useful in gene therapy  
CC to treat a clinical disorder, to study RNA processing and/or gene  
CC regulation. The present sequence represents an uncorrected version of  
CC beta-domain deleted factor VIII protein encoding gene sequence (construct  
CC pcy-2). This was used to develop a new coding sequence for beta-domain  
CC deleted factor VIII protein by correcting the consensus splice sites.  
XX  
XX  
SQ Sequence 9164 BP; 2451 A; 2170 C; 2081 G; 2462 T; 0 other;

Query Match 69.9%; Score 5553; DB 20; Length 9164;  
Best Local Similarity 84.7%; Pred. No. 0;  
Matches 6800; Conservative 0; Mismatches 130; Indels 1099; Gaps 7;

QY 415 ccacatgcaaatagagctccacacgtctcttcgtgcctttcgagatctcttta 474  
DB 1001 ccacatggaatagagctccacacgtctcttcgtgcctttcgagatctcttta 1060  
QY 475 gtgcacacagaagatatacactctgggtgacgtggaactgtcatggagacatatgcaagt 534  
DB 1061 gtgcacacagaagatatacactctgggtgacgtggaactgtcatggagacatatgcaagt 1120  
QY 535 atctcgtgtgagctgcctgtgacgcaagaattcctcctagagtgccaaatctttccat 594  
DB 1121 atctcgtgtgagctgcctgtgacgcaagaattcctcctagagtgccaaatctttccat 1180  
QY 595 tcaaacactcagctgtgtacaaaaaagactcgttgtagaattccggttaaccctttca 654  
DB 1181 tcaaacactcagctgtgtacaaaaaagactcgttgtagaattccggttaaccctttca 1240  
QY 655 acatcgttaagccaagccacactggaatggttctgtagtacctacatccagcgttag 714  
DB 1241 acatcgttaagccaagccacactggaatggttctgtagtacctacatccagcgttag 1300  
QY 715 ttatagatcacgtgtgcaacttaagaacatggtctccatccctgtcagttccatg 774  
DB 1301 ttatagatcacgtgtgcaacttaagaacatggtctccatccctgtcagttccatg 1360  
QY 775 ctgtgtgtatctactactggaagcttctgagggagctggaatatgatatacagaccatc 834  
DB 1361 ctgtgtgtatctactactggaagcttctgagggagctggaatatgatatacagaccatc 1420  
QY 835 aaaggggaaagaagatgataaagtcttcctctgtgtgaaagcatatattctggcagg 894  
DB 1421 aaaggggaaagaagatgataaagtcttcctctgtgtgaaagcatatattctggcagg 1480  
QY 895 tccgtgaagaagatggtccaaatggcctctgacccaactgtgcttaactactatacttt 954  
DB 1481 tccgtgaagaagatggtccaaatggcctctgacccaactgtgcttaactactatacttt 1540  
QY 955 ctcatgtgagactcgttaaaagactgaaattcagagcctcatgtgagcctactagatgta 1014  
DB 2621 ctcatgtgagactcgttaaaagactgaaattcagagcctcatgtgagcctactagatgta 1060

DB 1541 ctcatgtgagactcgttaaaagactgaaattcagagcctcatgtgagcctactagatgta 1600  
QY 1015 gagaaggagctctggccaaggaagacacagactctgacaaattatactacttttg 1074  
DB 1601 gagaaggagctctggccaaggaagacacagactctgacaaattatactacttttg 1660  
QY 1075 ctgtattgtatgaaaggaaagtgtggcaactcagaacaaagaactcctgtgacagata 1134  
DB 1661 ctgtattgtatgaaaggaaagtgtggcaactcagaacaaagaactcctgtgacagata 1720  
QY 1135 gggatgctgcatctgtcctgggccttgcttaaaatgacacacgtcaatggttatgtaaca 1194  
DB 1721 gggatgctgcatctgtcctgggccttgcttaaaatgacacacgtcaatggttatgtaaca 1780  
QY 1195 gctcctcgaagctgtgattgattgattcccaagaaatcagttatgtgattgtgtaa 1254  
DB 1781 gctcctcgaagctgtgattgattgattcccaagaaatcagttatgtgattgtgtaa 1840  
QY 1255 tgggcaacactctgaaagtgcactcaaatatcctctogaagtgacacattctgtgaga 1314  
DB 1841 tgggcaacactctgaaagtgcactcaaatatcctctogaagtgacacattctgtgaga 1900  
QY 1315 accatcgcacagcgctccttggaatctgcgaataactcctctactgctcaaacactct 1374  
DB 1901 accatcgcacagcgctccttggaatctgcgaataactcctctactgctcaaacactct 1960  
QY 1375 tgatgaccttggaacgtttcactactgtttgtcatatctctccaccaacatgatgca 1434  
DB 1961 tgatgaccttggaacgtttcactactgtttgtcatatctctccaccaacatgatgca 2020  
QY 1435 tggaaagctatgtcaaaagtagacagctgtccagaagaaacccaactaagaaatgaaata 1494  
DB 2021 tggaaagctatgtcaaaagtagacagctgtccagaagaaacccaactaagaaatgaaata 2080  
QY 1495 atgaaagacgggaagacatgatgatgtcttactgattctgaaatgagtgtgcaagt 1554  
DB 2081 atgaaagacgggaagacatgatgatgtcttactgattctgaaatgagtgtgcaagt 2140  
QY 1555 ttgatgtgacaacactcctctccttatccaatctgctcagttgccaagaagacactca 1614  
DB 2141 ttgatgtgacaacactcctctccttatccaatctgctcagttgccaagaagacactca 2200  
QY 1615 aaacttgggtacatcaactcgtcctcgaagagagagctggaactatgctcccttagtc 1674  
DB 2201 aaacttgggtacatcaactcgtcctcgaagagagagctggaactatgctcccttagtc 2260  
QY 1675 tcgcccgcgatgacgaagatgataaaagtcaatttggaaatggtccccaagcgagat 1734  
DB 2261 tcgcccgcgatgacgaagatgataaaagtcaatttggaaatggtccccaagcgagat 2320  
QY 1735 gtaggaagtaacaaaaagctccgatttatggtacatacagaatgaaacctttaagaactcgt 1794  
DB 2321 gtaggaagtaacaaaaagctccgatttatggtacatacagaatgaaacctttaagaactcgt 2380  
QY 1795 aagctatcagcatgaatcaggaatcttgggaacttaactttaagggaagttggagaca 1854  
DB 2381 aagctatcagcatgaatcaggaatcttgggaacttaactttaagggaagttggagaca 2440  
QY 1855 cactgttatattatttaagaatacaagcaagcagacatataacttaccctcaaggaa 1914  
DB 2441 cactgttatattatttaagaatacaagcaagcagacatataacttaccctcaaggaa 2500  
QY 1915 tcaactggtccgtcccttggatccaaggaatattccaagaagtgtaaacatttgagg 1974  
DB 2501 tcaactggtccgtcccttggatccaaggaatattccaagaagtgtaaacatttgagg 2560  
QY 1975 attttccaattctgcaggagaaatatccaataataatgagacatgactgtgaaagt 2034  
DB 2561 attttccaattctgcaggagaaatatccaataataatgagacatgactgtgaaagt 2620  
QY 2035 ggcacaactaaatcagatcctcgtgtgcctgacccgctattactcagtttgcgttaaatg 2094  
DB 2621 ggcacaactaaatcagatcctcgtgtgcctgacccgctattactcagtttgcgttaaatg 2680



QY 2095 agagagatcagcttcaggactcatctggccctccctcatctctgtcttaacaagaatctgag 2154  
|||||  
Db 2681 agagagatcagcttcaggactcatctggccctccctcatctctgtcttaacaagaatctgag 2740  
QY 2155 atcaaaagagaaacccaataatgtccagaagaagaatgtccatctgtttctgtatttg 2214  
|||||  
Db 2741 atcaaaagagaaacccaataatgtccagaagaagaatgtccatctgtttctgtatttg 2800  
QY 2215 atggagaacagagctgtgacctccacagagaaatatacaagcttctccccaatccagctg 2274  
|||||  
Db 2801 atggagaacagagctgtgacctccacagagaaatatacaagcttctccccaatccagctg 2860  
QY 2275 gaggcagctgagagatcccaagctccaagctccaacatcatgacaagcaatcaatgct 2334  
|||||  
Db 2861 gaggcagctgagagatcccaagctccaagctccaacatcatgacaagcaatcaatgct 2920  
QY 2335 atgttttgatagtttgcaagttgtcagttgtttgtcataggtggcatactgttcatctc 2394  
Db 2921 atgttttgatagtttgcaagttgtcagttgtttgtcataggtggcatactgttcatctc 2980  
QY 2395 taagcatgtgggacacagctgactctctctctctctctctctctcgatataccttcaaac 2454  
|||||  
Db 2981 taagcatgtgggacacagctgactctctctctctctctctctctcgatataccttcaaac 3040  
QY 2455 acaaaatgctctatgaaagacacactcaacctatcccatctccagggagaaatcttctca 2514  
|||||  
Db 3041 acaaaatgctctatgaaagacacactcaacctatcccatctccagggagaaatcttctca 3100  
QY 2515 tgtcgtatggaaaaacccaggtctatgatatctgtgggtgccaacactcagacttctcgaaaca 2574  
Db 3101 tgtcgtatggaaaaacccaggtctatgatatctgtgggtgccaacactcagacttctcgaaaca 3160  
QY 2575 gaggcagcagccgcttctctgaagtttctgtgtgtgacagaagaacacgtgtattatcag 2634  
|||||  
Db 3161 gaggcagcagccgcttctctgaagtttctgtgtgtgacagaagaacacgtgtattatcag 3220  
QY 2635 aggcagcttatgaaagatattcagacatactgtcgtgaataaacaatgccaattgaaccaa 2694  
|||||  
Db 3221 aggcagcttatgaaagatattcagacatactgtcgtgaataaacaatgccaattgaaccaa 3280  
QY 2695 gaagcttctcccaagaattcaagacacctagcactagcaaaagaaatltaatgccacc 2754  
|||||  
Db 3281 gaagcttctcccaaga-----cc 3298  
QY 2755 caccagcttgaaaacgcgcatacaagggaaataactgttactaccttccagtcagaataag 2814  
|||||  
Db 3299 caccagcttgaaaacgcgcatacaagggaaataactgttactaccttccagtcagaataag 3358  
QY 2815 aggaatctgactatgatatcaatcaagttgaattgaagaagaagaatttgaactt 2874  
|||||  
Db 3359 aggaatctgactatgatatcaatcaagttgaattgaagaagaagaatttgaactt 3418  
QY 2875 atgataagagatgaaaatccagagccccgcagcttccaagaagaaacacagacataattta 2934  
|||||  
Db 3419 atgataagagatgaaaatccagagccccgcagcttccaagaagaaacacagacataattta 3478  
QY 2935 ttgtctgcaagtggagaagctctcgtggatataatggatagctccccaatgctttaaagaa 2994  
|||||  
Db 3479 ttgtctgcaagtggagaagctctcgtggatataatggatagctccccaatgctttaaagaa 3538  
QY 2995 acaagggtccagagctgagcagtgctccctcagttccaagaagaatgttttccaagaattttag 3054  
|||||  
Db 3539 acaagggtccagagctgagcagtgctccctcagttccaagaagaatgttttccaagaattttag 3558  
QY 3055 atggctccttactcagcccttataccgttggagaactaaatgaaacatttggagctcttg 3114  
|||||  
Db 3599 atggctccttactcagcccttataccgttggagaactaaatgaaacatttggagctcttg 3658  
QY 3115 ggcacatatataagcaggaagtgtgaagataataatatacagtgaacttcaagaatcaggcct 3174  
|||||  
Db 3659 ggcacatatataagcaggaagtgtgaagataataatatacagtgaacttcaagaatcaggcct 3718

QY 3175 ctgcctccatctctctatctatctagccttatcttcttaagaaagatccagggcaagag 3234  
|||||  
Db 3719 ctgcctccatctctctatctatctagccttatcttcttaagaaagatccagggcaagag 3778  
QY 3235 cagaacctagaaaaaacttgttcaagcccaatgaaacccaaacttacttcttgaaagtgc 3294  
|||||  
Db 3779 cagaacctagaaaaaacttgttcaagcccaatgaaacccaaacttacttcttgaaagtgc 3838  
QY 3295 aacatcatatggccaccactaaagatgattttagcttgcagaagcttggctatttctctg 3354  
Db 3839 aacatcatatggccaccactaaagatgattttagcttgcagaagcttggctatttctctg 3898  
QY 3355 atgttgaaccttgaaaaagatgtgacacccaagctgatttggacccctctgtctgcaca 3414  
Db 3899 atgttgaaccttgaaaaagatgtgacacccaagctgatttggacccctctgtctgcaca 3958  
QY 3415 ctaacacacttgaaacctgtctcatgtggagacaagtgaacagttaagaaattgtctgtttt 3474  
Db 3959 ctaacacacttgaaacctgtctcatgtggagacaagtgaacagttaagaaattgtctgtttt 4018  
QY 3475 tcaccatcttcttgatgaaacccaagctggttactctcacttgaaataatgtgaaagaactgca 3534  
|||||  
Db 4019 tcaccatcttcttgatgaaacccaagctggttactctcacttgaaataatgtgaaagaactgca 4078  
QY 3535 ggcctcccttgacatataccaagtgaagatcccaactttaaagaattatcgcttccatg 3594  
Db 4079 ggcctcccttgacatataccaagtgaagatcccaactttaaagaattatcgcttccatg 4138  
QY 3595 caatcaatgtgctataataatgatacactacccgtgttagtaatgtgtcagatccaagga 3654  
Db 4139 caatcaatgtgctataataatgatacactacccgtgttagtaatgtgtcagatccaagga 4198  
QY 3655 ttcgatgttatctgtcagacatggcgacgaatggaacacatcatctattcatcttgctg 3714  
Db 4199 ttcgatgttatctgtcagacatggcgacgaatggaacacatcatctattcatcttgctg 4258  
QY 3715 gacatgtcttcactgtacgaaaaaagagagataaaatgtgacactgtacacatctatc 3774  
Db 4259 gacatgtcttcactgtacgaaaaaagagagataaaatgtgacactgtacacatctatc 4318  
QY 3775 caagtgttttcttgagaacgttggaatgtttacaatccaaagcttgaaatttggcgggtggaat 3834  
|||||  
Db 4319 caagtgttttcttgagaacgttggaatgtttacaatccaaagcttgaaatttggcgggtggaat 4378  
QY 3835 gacctatttggcagacatctcatgctgtggatgagcaaaccttctctgttggacagcata 3894  
Db 4379 gacctatttggcagacatctcatgctgtggatgagcaaaccttctctgttggacagcata 4438  
QY 3895 agtctcagactccctctgggaatgtctctggaacacatataagatttcagatlcagact 3954  
Db 4439 agtctcagactccctctgggaatgtctctggaacacatataagatttcagatlcagact 4498  
QY 3955 caaggacatatggaacagtgggcccccaagcttggccagactcatattctcggatcaatca 4014  
|||||  
Db 4499 caaggacatatggaacagtgggcccccaagcttggccagactcatattctcggatcaatca 4558  
QY 4015 atgccttgagacccaagagaccccttctcttgatcaaggttggatctgttggcaccatgaa 4074  
Db 4559 atgccttgagacccaagagaccccttctcttgatcaaggttggatctgttggcaccatgaa 4618  
QY 4075 ttattacaggcatcaagaacccaaggttgcctgccaagaagttctccagctcttaacatctc 4134  
Db 4619 ttattacaggcatcaagaacccaaggttgcctgccaagaagttctccagctcttaacatctc 4678  
QY 4135 agttatcatcatgtatagctcttgatggagaagaatgtgcagacttalcgaaggaaattcca 4194  
Db 4679 agttatcatcatgtatagctcttgatggagaagaatgtgcagacttalcgaaggaaattcca 4738  
QY 4195 ctggaaaccttaatgtgcttcttggcgaatgtgatcatctctgggataaaacacataattc 4254  
Db 4739 ctggaaaccttaatgtgcttcttggcgaatgtgatcatctctgggataaaacacataattc 4798  
QY 4255 ttaacctccaattattgtctcgatataatccgttltggaaacccaactcatattagatctgca 4314



[illegible][illegible]

|    |      |   |       |
|----|------|---|-------|
| QY | 5418 | ccagcgcttcccccctggaagccctccgtgtgcgtctccctgtgtccgaaccctgcgcgttac     | 54177 |
| Db | 7019 | ccagcgcttcccccctggaagccctccgtgtgcgtctccctgtgtccgaaccctgcgcgttac     | 7078  |
| QY | 5478 | cgagtaacctgcccgccttccctccctccgggaagcgtgagcgctcttccatgctcaagctg      | 5537  |
| Db | 7079 | cgagtaacctgcccgccttccctccctccgggaagcgtgagcgctcttccatgctcaagctg      | 7138  |
| QY | 5538 | taagtatctcaagttcgtgtgaagtcgtctccgaagcggcgctgtgtgtgcagaaacccc        | 5597  |
| Db | 7139 | taagtatctcaagttcgtgtgaagtcgtctccgaagcggcgctgtgtgtgcagaaacccc        | 7198  |
| QY | 5598 | cggttcagcccggaacgcgcgcgtccctaccggttaactatcgctcttgagttccaaaccggttaag | 5657  |
| Db | 7139 | cggttcagcccggaacgcgcgcgtccctaccggttaactatcgctcttgagttccaaaccggttaag | 7258  |
| QY | 5658 | acacgactatcgcgcactgtgcagccagccactgtgtlaacagatctagcagagcgagttagt     | 5717  |
| Db | 7259 | acacgactatcgcgcactgtgcagccagccactgtgtlaacagatctagcagagcgagttagt     | 7318  |
| QY | 5718 | agcgcggtgtcctacagagttctctgaagtgtgtgtgccttaactacggtcactagaaagacgt    | 5777  |
| Db | 7319 | agcgcggtgtcctacagagttctctgaagtgtgtgtgccttaactacggtcactagaaagacgt    | 7378  |
| QY | 5778 | attcggtatctgcgcctctgtgtgaagccaattacaccttcggaaaaagagttgttgactctg     | 5837  |
| Db | 7379 | attcggtatctgcgcctctgtgtgaagccaattacaccttcggaaaaagagttgttgactctg     | 7438  |
| QY | 5838 | atccgcgcaaaacaaacccacgcgtgtgtagcgtgtcttctgtgttgcaagcagcagatbac      | 5897  |
| Db | 7439 | atccgcgcaaaacaaacccacgcgtgtgtagcgtgtcttctgtgttgcaagcagcagatbac      | 7498  |
| QY | 5898 | gcgcaagaaaaaaggatctctcaagaagaatcccttgatctcttcttaaggggtctgcagctca    | 5957  |
| Db | 7499 | gcgcaagaaaaaaggatctctcaagaagaatcccttgatctcttcttaaggggtctgcagctca    | 7558  |
| QY | 5958 | gttgaaacgaaacacccacgctgttaaagggtatttgtgtcaltgagatattacaagaagatcttac | 6017  |
| Db | 7559 | gttgaaacgaaacacccacgctgttaaagggtatttgtgtcaltgagatattacaagaagatcttac | 7618  |
| QY | 6018 | ctagaacctttaaattaaatgaagagttttaaalcactaaagtatatatgtgaataac          | 6077  |
| Db | 7619 | ctagaacctttaaattaaatgaagagttttaaalcactaaagtatatatgtgaataac          | 7678  |
| QY | 6078 | ttgtgtcgcagagttacaacatgagttactaagtgtagagcaactctcagcgatctcttaatt     | 6137  |
| Db | 7679 | ttgtgtcgcagagttacaacatgagttactaagtgtagagcaactctcagcgatctcttaatt     | 7738  |
| QY | 6138 | tcgttcatcatalagtgtgcgtgactcccggtcgtgtgtagaataacatgatacggagggcgt     | 6197  |
| Db | 7739 | tcgttcatcatalagtgtgcgtgactcccggtcgtgtgtagaataacatgatacggagggcgt     | 7798  |
| QY | 6198 | accatctgcgcccagcgtgtgcgaatgtagtaccggagagaccacgcgtctacacgggtccagattt | 6257  |
| Db | 7799 | accatctgcgcccagcgtgtgcgaatgtagtaccggagagaccacgcgtctacacgggtccagattt | 7858  |
| QY | 6258 | atcagcagataaacccagccagcggaaaggcgcgagagcgcagaaagtgtgtccctgcgaacttacc | 6317  |
| Db | 7859 | atcagcagataaacccagccagcggaaaggcgcgagagcgcagaaagtgtgtccctgcgaacttacc | 7918  |
| QY | 6318 | cggtcccatccagttcatattatgtgtgtgcggggaagctagagttaagttctgcgaagttaa     | 6377  |
| Db | 7919 | cggtcccatccagttcatattatgtgtgtgcggggaagctagagttaagttctgcgaagttaa     | 7978  |
| QY | 6378 | taagtttgcgcaagttgtgttccatctgtcaacagatctcgtgtgtgtcaagcgtctgtctgttg   | 6437  |
| Db | 7979 | taagtttgcgcaagttgtgttccatctgtcaacagatctcgtgtgtgtcaagcgtctgtctgttg   | 8038  |
| QY | 6438 | tatgtgtcatcgaagtcgcgtgtcccaacgatacgaagcggagttacatgataccccaagt       | 6497  |
| Db | 8039 | tatgtgtcatcgaagtcgcgtgtcccaacgatacgaagcggagttacatgataccccaagt       | 8098  |

|    |      |  |      |
|----|------|--|------|
| QY | 6438 | gtgcacaaaagcgggtatgcctccttcggtctctccgatacgtgtccagaagtaagtatggccgc    | 6557 |
| Db | 8099 | gtgcacaaaagcgggtatgcctccttcggtctctccgatacgtgtccagaagtaagtatggccgc    | 8158 |
| QY | 6558 | agtgtatcatcatctgtgttatggcaagcgtacatactctactgcatatccatcgt             | 6617 |
| Db | 8159 | agtgtatcatcatctgtgttatggcaagcgtacatactctactgcatatccatcgt             | 8218 |
| QY | 6618 | aagatgctttctctgtgacgtgtgagtaactaaccaagatattctatgatagtatgcg           | 6677 |
| Db | 8219 | aagatgctttctctgtgacgtgtgagtaactaaccaagatattctatgatagtatgcg           | 8278 |
| QY | 6678 | gcgcacgcgagtggtcctcttcggccgcgcgtcaaacggtatactacgcgcgcacataagacac     | 6737 |
| Db | 8279 | gcgcacgcgagtggtcctcttcggccgcgcgtcaaacggtatactacgcgcgcacataagacac     | 8338 |
| QY | 6738 | ttcaaaagtgtctcatctatgtgaaaaagcttctcttcgggcgcaaaactctcaagatcttacc     | 6797 |
| Db | 8339 | ttcaaaagtgtctcatctatgtgaaaaagcttctcttcgggcgcaaaactctcaagatcttacc     | 8398 |
| QY | 6798 | gctgtttgagatccagcttcctcgatgtgaataccacatcgtgacccaactgatatcttccagatctt | 6857 |
| Db | 8399 | gctgtttgagatccagcttcctcgatgtgaataccacatcgtgacccaactgatatcttccagatctt | 8458 |
| QY | 6858 | tactttccaccgcgtttctctggtgtgagcaaaaacaggaagcgaatagtccgcgcaaaaagcg     | 6917 |
| Db | 8459 | tactttccaccgcgtttctctggtgtgagcaaaaacaggaagcgaatagtccgcgcaaaaagcg     | 8518 |
| QY | 6918 | aataagcgcgacacgcgaatagtttgaaatactatactctctcttccaatataatltgaag        | 6977 |
| Db | 8519 | aataagcgcgacacgcgaatagtttgaaatactatactctctcttccaatataatltgaag        | 8578 |
| QY | 6978 | cattatcaaggttatatgttctcatctggcgagatatactatttgataattttagaaaaataa      | 7037 |
| Db | 8579 | cattatcaaggttatatgttctcatctggcgagatatactatttgataattttagaaaaataa      | 8638 |
| QY | 7038 | acaaataaggggtttccgcgcacatcttcccgcaaaaagtgcaacccttgacgtctgaagaaacat   | 7097 |
| Db | 8639 | acaaataaggggtttccgcgcacatcttcccgcaaaaagtgcaacccttgacgtctgaagaaacat   | 8698 |
| QY | 7098 | tatatcatatgcatactataacctaataaataatggcgtatatcaagagcccttctgtctccgcg    | 7157 |
| Db | 8699 | tatatcatatgcatactataacctaataaataatggcgtatatcaagagcccttctgtctccgcg    | 8758 |
| QY | 7158 | tttcggtgatatacgtgtaaaaactctcgacacatctgcagcgtcccgagagacggttcaacgctg   | 7217 |
| Db | 8759 | tttcggtgatatacgtgtaaaaactctcgacacatctgcagcgtcccgagagacggttcaacgctg   | 8818 |
| QY | 7218 | tctgtgaagcggatgctgcgggagcagacaagcccgatcaaggcgcgttcaagcgggtgttgcgcg   | 7277 |
| Db | 8819 | tctgtgaagcggatgctgcgggagcagacaagcccgatcaaggcgcgttcaagcgggtgttgcgcg   | 8878 |
| QY | 7278 | gtgtccggcggcgtgcgttttaactatgcgcatacagaagcagatgtttacgtgaagtgacacatat  | 7337 |
| Db | 8879 | gtgtccggcggcgtgcgttttaactatgcgcatacagaagcagatgtttacgtgaagtgacacatat  | 8938 |
| QY | 7338 | gcgggtgtgaataaccgcacagatgcgttlaaggagaataataccgcatacagg               | 7386 |
| Db | 8939 | gcgggtgtgaataaccgcacagatgcgttlaaggagaataataccgcatacagg               | 8987 |

RESULT 3  
AAD00122  
ID AAD00122 standard; DNA; 4999 BP.  
XX  
XX AAD00122;  
AC  
XX  
XX  
DT 31-JUL-2000 (first entry)  
XX  
XX  
DE Recombinant adeno associated vector construct, pVn4.lcF8deltaB.  
XX  
XX Recombinant Adeno Associated Vector; rAAV; pVn4.lcF8deltaB; promoter  
KW human Factor VIII; hFVIII; EF1alpha; human elongation factor-1alpha;



|    |      |   |      |
|----|------|---|------|
| QY | 1393 | ttctactgttttgcatactcttccccaacatgcatggaagcttaatgcaag               | 1452 |
| Db | 1389 | ttctactgttttgcatactcttccccaacatgcatggaagcttaatgcaag               | 1448 |
| QY | 1453 | tagacagctgtccagaggaaacccaactacgaatgaaaaataatgaaagcggaaagct        | 1512 |
| Db | 1449 | tagacagctgtccagaggaaacccaactacgaatgaaaaataatgaaagcggaaagct        | 1508 |
| QY | 1513 | atgataagatcttactgattcttgaatggaatggtgtcaggttttgatgatacaactctc      | 1572 |
| Db | 1509 | atgataagatcttactgattcttgaatggaatggtgtcaggttttgatgatacaactctc      | 1568 |
| QY | 1573 | cttcctttatccaatctgcgtcagttgtccaagaagcatcccaaaacttgggttacatcca     | 1632 |
| Db | 1569 | cttcctttatccaatctgcgtcagttgtccaagaagcatcccaaaacttgggttacatcca     | 1628 |
| QY | 1633 | tttgtgtcgaagaggaggaactcgtggaactgctcccttaagttccctcgccccgaatgacaga  | 1692 |
| Db | 1629 | tttgtgtcgaagaggaggaactcgtggaactgctcccttaagttccctcgccccgaatgacaga  | 1688 |
| QY | 1693 | gttataaagatcaatatatttgacaatggtccttcagcggtatgtgttagagatgatacaaaaag | 1752 |
| Db | 1689 | gttataaagatcaatatatttgacaatggtccttcagcggtatgtgttagagatgatacaaaaag | 1748 |
| QY | 1753 | tcagattatgcatcacacagatgaacctttaagactcgtgaagctatctacgaatggaat      | 1812 |
| Db | 1749 | tcagattatgcatcacacagatgaacctttaagactcgtgaagctatctacgaatggaat      | 1808 |
| QY | 1813 | caggaaatcttgggaacttacttatagtgggaagtgtgagacacactgttgaattatatta     | 1872 |
| Db | 1809 | caggaaatcttgggaacttacttatagtgggaagtgtgagacacactgttgaattatatta     | 1868 |
| QY | 1873 | agaatcaagcagaagacacataatacatctacccctcaaggaatacaatgatactcgctctt    | 1932 |
| Db | 1869 | agaatcaagcagaagacacataatacatctacccctcaaggaatacaatgatactcgctctt    | 1928 |
| QY | 1933 | tgatccaagagaattacccaagaagtgtgaacaactttgaagaattttccaattccgcag      | 1992 |
| Db | 1929 | tgatccaagagaattacccaagaagtgtgaacaactttgaagaattttccaattccgcag      | 1988 |
| QY | 1993 | gagaataattcaaatataaatgtgacaagtgtactgtagaagaatgggccaactaaatcaatc   | 2052 |
| Db | 1989 | gagaataattcaaatataaatgtgacaagtgtactgtagaagaatgggccaactaaatcaatc   | 2048 |
| QY | 2053 | ctcgggtgcgtgacccggtattactctagtttcgttaatatgtagaagagatcttaactcag    | 2112 |
| Db | 2049 | ctcgggtgcgtgacccggtattactctagtttcgttaatatgtagaagagatcttaactcag    | 2108 |
| QY | 2113 | gactcatgtgcccctctcctcactcgtctacaagaatctgtatgatacgaagaggaaccaga    | 2172 |
| Db | 2109 | gactcatgtgcccctctcctcactcgtctacaagaatctgtatgatacgaagaggaaccaga    | 2168 |
| QY | 2173 | taatgtcagacaagaagaatgtcatcctgttttctgtatattgtatgtagaaccggaactgtc   | 2232 |
| Db | 2169 | taatgtcagacaagaagaatgtcatcctgttttctgtatattgtatgtagaaccggaactgtc   | 2228 |
| QY | 2233 | accctcaagaagaataatacaagcttcttccccaatccaagctcggtgtcagctttagagatc   | 2292 |
| Db | 2229 | accctcaagaagaataatacaagcttcttccccaatccaagctcggtgtcagctttagagatc   | 2288 |
| QY | 2293 | cagaagctccaagcctccaacaatcagcagaacataatgtgtaagtttttatagtttgc       | 2352 |
| Db | 2289 | cagaagctccaagcctccaacaatcagcagaacataatgtgtaagtttttatagtttgc       | 2348 |
| QY | 2353 | agttgtcagtttgttctgcatgaggtagcatatcgtgtacatcttaagcatvtggagcacaga   | 2412 |
| Db | 2349 | agttgtcagtttgttctgcatgaggtagcatatcgtgtacatcttaagcatvtggagcacaga   | 2408 |
| QY | 2413 | ctgactcctcttctgtcttcttctctgtagatacttccaacaacaaaatggtctataag       | 2472 |
| Db | 2409 | ctgactcctcttctgtcttcttctctgtagatacttccaacaacaaaatggtctataag       | 2468 |
| QY | 2473 | acacactcaacctatctcccatctccagagagaactgtctcatgtcgaatgaaaaaccag      | 2532 |

---

|    |      |  |      |
|----|------|--|------|
| Db | 2469 | aaacactaaccttatctccatcttccaggaagaactgtcttcatgtcaggtgaaaaaccag      | 2528 |
| QY | 2533 | gtctatggatctctgggtgtgcacaactcagactctcggaaacagagcatgacgccttac       | 2592 |
| Db | 2529 | gtctatggatctctgggtgtgcacaactcagactctcggaaacagagcatgacgccttac       | 2588 |
| QY | 2593 | tgaaggttcttagttgtgtacagaagacactggtgtgattatacagaagacagttatgaagta    | 2652 |
| Db | 2589 | tgaaggttcttagttgtgtacagaagacactggtgtgattatacagaagacagttatgaagta    | 2648 |
| QY | 2653 | tttcagcatcttctgtgtgtaaaaaacaatgcacatgtaaccagaagctctccccaagt        | 2712 |
| Db | 2649 | tttcagcatcttctgtgtgtaaaaaacaatgcacatgtaaccagaagctctccccaagt        | 2707 |
| QY | 2713 | caagacacccctagacactagcgcaaaagcaatttaatgtccaccaccagttctgaaagcc      | 2772 |
| Db | 2708 | caagacacccctagacactagcgcaaaagcaatttaatgtccaccaccagttctgaaagcc      | 2726 |
| QY | 2773 | atcaacgggaaataactcgttactcttcagtcagatcaagagaatgtgataatg             | 2832 |
| Db | 2727 | atcaacgggaaataactcgttactcttcagtcagatcaagagaatgtgataatg             | 2786 |
| QY | 2833 | ataccatactcagttgaaatgaaagagaatcttgacattatgataggaatgaaatc           | 2892 |
| Db | 2787 | ataccatactcagttgaaatgaaagagaatcttgacattatgataggaatgaaatc           | 2846 |
| QY | 2893 | agagcccccagacttccaagaagaacaacagacataattatgtcgtcagttgagagagc        | 2952 |
| Db | 2847 | agagcccccagacttccaagaagaacaacagacataattatgtcgtcagttgagagagc        | 2906 |
| QY | 2953 | ttctggatattgggaatgtagtgcctccacatggttctaagaacaaggtctcaagatgtgca     | 3012 |
| Db | 2907 | ttctggatattgggaatgtagtgcctccacatggttctaagaacaaggtctcaagatgtgca     | 2966 |
| QY | 3013 | gtgtcccccagttcaagaagaattgtttccaggaatttaacgtatggtcccttaactcagc      | 3072 |
| Db | 2967 | gtgtcccccagttcaagaagaattgtttccaggaatttaacgtatggtcccttaactcagc      | 3026 |
| QY | 3073 | ccctataccgttggagaaactaaatgaaacatcttggaatccctcgtgggcacataataagagcag | 3132 |
| Db | 3027 | ccctataccgttggagaaactaaatgaaacatcttggaatccctcgtgggcacataataagagcag | 3086 |
| QY | 3133 | aagttgaagaataatactatgtaactttcagaagaactcagcctctcgtccctatctctct      | 3192 |
| Db | 3087 | aagttgaagaataatactatgtaactttcagaagaactcagcctctcgtccctatctctct      | 3146 |
| QY | 3193 | attcctagccttattcttataagagaatcagagagcagaagcaagaactgaaaaaact         | 3252 |
| Db | 3147 | attcctagccttattcttataagagaatcagagagcagaagcaagaactgaaaaaact         | 3206 |
| QY | 3253 | ttgtcaagcctaattgaaacccaacttacttctggaagatgtcaacatcatatgtgaccca      | 3312 |
| Db | 3207 | ttgtcaagcctaattgaaacccaacttacttctggaagatgtcaacatcatatgtgaccca      | 3266 |
| QY | 3313 | ctaaagatgaatttgactgcgcaaaagccttgggtttatcttcgtatggttgaacttgaaaaag   | 3372 |
| Db | 3267 | ctaaagatgaatttgactgcgcaaaagccttgggtttatcttcgtatggttgaacttgaaaaag   | 3326 |
| QY | 3373 | atgtgacactcagcctgtaattggaacccctctcgtctcgtccaacataacacatgtgaacctg   | 3432 |
| Db | 3327 | atgtgacactcagcctgtaattggaacccctctcgtctcgtccaacataacacatgtgaacctg   | 3386 |
| QY | 3433 | ctcatvggagagaagatgacagatagcaagaattgtcgtgttttccaacatcttgaatgaga     | 3492 |
| Db | 3387 | ctcatvggagagaagatgacagatagcaagaattgtcgtgttttccaacatcttgaatgaga     | 3446 |
| QY | 3493 | ccaaaagctgtaatttacttaacttaaaataatgaaaaaactgtgaaggtcccccgaatctcc    | 3552 |
| Db | 3447 | ccaaaagctgtaatttacttcaactgaaaaataatgaaaaaactgtgaaggtcccccgaatctcc  | 3506 |
| QY | 3553 | agatggaagatcccaactttaaagagaattatcgtctccaatgcaatggtcacataa          | 3612 |

```
Db 3507 agatggaagatcccaatttlaaagaagatlaacgtctccatgcaatcaatggctacataa 3566
QY 3613 tggatatacactcctgcttaagtaagctcagaatcaaaagatctgatgtctctca 3672
Db 3567 tggatacactcactcctgcttaagtaagctcagaatcaaaagatctgatgtctctca 3626
QY 3673 gcatggtgagcaatgaaaaacatcctatcatctcagtgagacatgtctcaactgac 3732
Db 3627 gcatggtgagcaatgaaaaacatcctatcatctcagtgagacatgtctcaactgac 3686
QY 3733 gaaaaaagaagagatataaataagcagctgataatctctcagtggtgttttgagacg 3792
Db 3687 gaaaaaagaagagatataaataagcagctgataatctctcagtggtgttttgagacg 3746
QY 3793 tggaaatcttaccatcccaagcttggaatttggcgggtggaagcttatctggcgaacac 3852
Db 3747 tggaaatcttaccatcccaagcttggaatttggcgggtggaagcttatctggcgaacac 3806
QY 3853 tacaatctggaatgagcaactttctctgtgtacagcaataagtgtcagactccctgg 3912
Db 3807 tacatgtcggatgagcaactttctctgtgtacagcaataagtgtcagactccctgg 3866
QY 3913 gaatggtcttcggaaacttgaagatttccagatacagcttcaggaacataatgacac 3972
Db 3867 gaatggtcttcggaaacttgaagatttccagatacagcttcaggaacataatgacac 3926
QY 3973 gggcccaaaagctggccagacttcatctccgaatcaatcagctcgtgagacacaaag 4032
Db 3927 gggcccaaaagctggccagacttcatctccgaatcaatcagctcgtgagacacaaag 3986
QY 4033 agcccttcttcttgatacaagtgatctgttggaaccaatgatatccacggcatcaaga 4092
Db 3987 agcccttcttcttgatacaagtgatctgttggaaccaatgatatccacggcatcaaga 4046
QY 4093 cccgaggtgcgcgtcagaagttctccagctctacatctcctgaattacatcatgtata 4152
Db 4047 cccgaggtgcgcgtcagaagttctccagctctacatctcctgaattacatcatgtata 4106
QY 4153 gtcttgatggaagaagtgagcaacttctcgagaaatccactggaaccttaatgtct 4212
Db 4107 gtcttgatggaagaagtgagcaacttctcgagaaatccactggaaccttaatgtct 4166
QY 4213 tcttggcaatgtggtatcatctcgtggaataaaacacataattttaaacctccaatatg 4272
Db 4167 tcttggcaatgtggtatcatctcgtggaataaaacacataattttaaacctccaatatg 4226
QY 4273 ctgcatatacgttcttgacccaactcttaagatctgcgcagcactcttcgcatggaat 4332
Db 4227 ctgcatatacgttcttgacccaactcttaagatctgcgcagcactcttcgcatggaat 4286
QY 4333 tgaatggtctgcatltaataatgttcgaacatgcatctgggaatggagaatlaagcaat 4392
Db 4287 tgaatggtctgcatltaataatgttcgaacatgcatctgggaatggagaatlaagcaat 4346
QY 4393 cagatgcacagatctactgtctcatctcacttaaccaatagtttgcacactgtctcct 4452
Db 4347 cagatgcacagatctactgtctcatctcacttaaccaatagtttgcacactgtctcct 4406
QY 4453 caaaagctcgactcactcactcaaaaggaggaatgacctggagactcccggaataaac 4512
Db 4407 caaaagctcgactcactcactcaaaaggaggaatgacctggagactcccggaataaac 4466
QY 4513 caaaaggtgtgctcaagtggaacttcagaaagacaaatgaaatccagagtaactactc 4572
Db 4467 caaaaggtgtgctcaagtggaacttcagaaagacaaatgaaatccagagtaactactc 4526
QY 4573 agggagtaaaatctcgtcttaccagcatgtatgtgaaggagttcctcatctccagac 4632
Db 4527 agggagtaaaatctcgtcttaccagcatgtatgtgaaggagttcctcatctccagac 4586
QY 4633 aagatggccatcagtggaactctcttctcgaatcggcaaaatlaaagttttcagggaa 4692
Db 4587 aagatggccatcagtggaactctcttctcgaatcggcaaaatlaaagttttcagggaa 4646
```

```
QY 4693 atcaagactccttcaaacactgtgtgtaactctctagacacacgcttactgactctacc 4752
Db 4647 atcaagactccttcaaacactgtgtgtaactctctctagacacacgcttactgactctacc 4706
QY 4753 ttcgaatcaaccccaagatgtgtgacacagaattgaccttgaagatgaggttctggct 4812
Db 4707 ttcgaatcaaccccaagatgtgtgacacagaattgaccttgaagatgaggttctggct 4766
QY 4813 ggcgaagcaaggaaccttactgactcgaagcgagttctctcgaaggagatcggcaataaa 4872
Db 4767 ggcgaagcaaggaaccttactgactcgaagcgagttctctcgaaggagatcggcaataaa 4805
QY 4873 gacagaataaaagcaagcgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4932
Db 4806 gaatttatcttcatctgaaatagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4865
QY 4933 gtgatggaattggcactcctctctgcgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 4992
Db 4866 gtgatggaattggcactcctctctgcgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 4925
QY 4993 aagcccgagcgtcggcgacacttggctgcgcggtcgaatgagcgagcgcgag 5051
Db 4926 aagtcgcgcgagcgccggcgttgcgcggcgtcgaatgagcgagcgcgag 4984
```

## RESULT 4

AAx88293 ID AAx88293 standard: DNA; 4629 BP.

AAx88293;

24-SEP-1999 (first entry)

Human Factor VIII with B domain deleted cDNA.

Adenoviral vector: Factor VIII; Factor IX; clotting factor: treatment:

haemostatic; haemophilia A; haemophilia B; gene therapy; ss.

Homo sapiens.

US5935935-A.

10-AUG-1999.

07-JUN-1995; 95US-0484891.

07-JUN-1995; 95US-0484891.

10-JUN-1993; 93US-0074920.

25-MAR-1994; 94US-0218335.

(GENE-) GENETIC THERAPY INC.

Connelly S, Kaleko M, Smith T;

WPI; 1999-457617/38.

Adenoviral vectors useful for treating hemophilia

Example 1; Column 53-58; 90pp; English.

This invention describes novel adenoviral vectors comprising at least

one DNA sequence encoding a clotting factor (Factor IX or Factor VIII).

The vectors of the invention have haemostatic activity. The vectors are

useful for the treatment of hemophilia A or hemophilia B by gene therapy.

This sequence represents human Factor VIII cDNA which has the B domain

deleted.

Sequence 4629 BP; 1319 A; 1050 C; 1015 G; 1245 T; 0 other;

Query Match 55.6%; Score 4413; DB 20; Length 4629;  
Best Local Similarity 99.8%; Pred. No. 0;

| Matches 4419; Conservative 0; Mismatches 10; Indels 0; Gaps 0; |      |  |      |  |
|--|------|--|------|--|
| QY   | 420  | atgcaaatagagctctccaccctgtcttctctctgctcttgagattctcttaagcc       | 479  |  |
| Db   | 1    | atgcaaatagagctctccaccctgtcttctctctgctcttgagattctcttaagcc       | 60   |  |
| QY   | 480  | accagaatatactacccctgggtgacagtgaactgtcacaggagatataatgatactc     | 539  |  |
| Db   | 61   | accagaatatactacccctgggtgacagtgaactgtcacaggagatataatgatactc     | 120  |  |
| QY   | 540  | ggtgagctgctgtgagcgaagaattctctctagagtgccaaatctttccatccaac       | 599  |  |
| Db   | 121  | ggtgagctgctgtgagcgaagaattctctctagagtgccaaatctttccatccaac       | 180  |  |
| QY   | 600  | accctgactgtgtaaaaaagactgtgtgtgaattccaggttccactttccaatc         | 659  |  |
| Db   | 181  | accctgactgtgtaaaaaagactgtgtgtgaattccaggttccactttccaatc         | 240  |  |
| QY   | 660  | gctaaagcgaagccacccttgatggtctgtagtctcaccatccagcgtgagttat        | 719  |  |
| Db   | 241  | gctaaagcgaagccacccttgatggtctgtagtctcaccatccagcgtgagttat        | 300  |  |
| QY   | 720  | gatacagtggtcattacacttaagaacatggtctccatccctgtaagttcagttcgt      | 779  |  |
| Db   | 301  | gatacagtggtcattacacttaagaacatggtctccatccctgtaagttcagttcgt      | 360  |  |
| QY   | 780  | ggtgtatccctactggaagagcttctgagggagctggaatgatatgatacagatccaag    | 839  |  |
| Db   | 361  | ggtgtatccctactggaagagcttctgagggagctggaatgatatgatacagatccaag    | 420  |  |
| QY   | 840  | gaagaagaatgatatgaatgaagctctccctggtggaagccatatagtctgagctctg     | 899  |  |
| Db   | 421  | gaagaagaatgatatgaatgaagctctccctggtggaagccatatagtctgagctctg     | 480  |  |
| QY   | 900  | aaagagaatgtccaatggtccctgagcccaatggtcctaactatacttctcat          | 939  |  |
| Db   | 481  | aaagagaatgtccaatggtccctgagcccaatggtcctaactatacttctcat          | 540  |  |
| QY   | 960  | gtggaactgtgtaaaaagactggaattcagagctatggaagccctactagattagaaga    | 1019 |  |
| Db   | 541  | gtggaactgtgtaaaaagactggaattcagagctatggaagccctactagattagaaga    | 600  |  |
| QY   | 1020 | gggagctggtcccaaggaagaagacagacctgtgcaacaatttatactactttgtcgt     | 1079 |  |
| Db   | 601  | gggagctggtcccaaggaagaagacagacctgtgcaacaatttatactactttgtcgt     | 660  |  |
| QY   | 1080 | tttgaatgaagggaaaagtgtgcaactcaagaacaaagactccttgatgaggaatggat    | 1139 |  |
| Db   | 661  | tttgaatgaagggaaaagtgtgcaactcaagaacaaagactccttgatgaggaatggat    | 720  |  |
| QY   | 1140 | gtcgcatactgtcgtgggctgtggtctaaatgacacagttcaatgtgtatgtaaacaggtct | 1199 |  |
| Db   | 721  | gtcgcatactgtcgtgggctgtggtctaaatgacacagttcaatgtgtatgtaaacaggtct | 780  |  |
| QY   | 1200 | ctgcagagttcgtatggtatggtccacaggaataatcagttcatgtgcatgtgtaatgggc  | 1259 |  |
| Db   | 781  | ctgcagagttcgtatggtatggtccacaggaataatcagttcatgtgcatgtgtaatgggc  | 840  |  |
| QY   | 1260 | accactctcgaagtgcataataatctcctggaagttcaacaaattctctgtggaagaacct  | 1319 |  |
| Db   | 841  | accactctcgaagtgcataataatctcctggaagttcaacaaattctctgtggaagaacct  | 900  |  |
| QY   | 1320 | cgcacagcgctcttggaatctcgcacaataacttctccttaactgtctcaaaaccccttgag | 1379 |  |
| Db   | 901  | cgcacagcgctcttggaatctcgcacaataacttctccttaactgtctcaaaaccccttgag | 960  |  |
| QY   | 1380 | gaacttggacagttcttactgttttgcataatctcttccacaacaatgatagtgcatgaa   | 1439 |  |
| Db   | 961  | gaacttggacagttcttactgttttgcataatctcttccacaacaatgatagtgcatgaa   | 1020 |  |
| QY   | 1440 | gcttatgtcaaaatagaacagctgtccaaggaagccccaactgaagatgaanaataatgaa  | 1499 |  |
| Db   | 1021 | gcttatgtcaaaatagaacagctgtccaaggaagccccaactgaagatgaanaataatgaa  | 1080 |  |

|    |      |  |      |  |
|----|------|--|------|--|
| QY | 1500 | gaagcggaaagactatgatatgattcttactgattctgnaatggatgtgctcaggtttgat  | 1559 |  |
| Db | 1081 | gaagcggaaagactatgatatgattcttactgattctgnaatggatgtgctcaggtttgat  | 1140 |  |
| QY | 1560 | gatgacaactctctctcccttatccaatctcgctcgaatttcgaagaagcatccaaact    | 1619 |  |
| Db | 1141 | gatgacaactctctctcccttatccaatctcgctcgaatttcgaagaagcatccaaact    | 1200 |  |
| QY | 1620 | tgggtacattacatgtctgtcgaaggaaggaactgggaactatgctcccttaagttccgc   | 1679 |  |
| Db | 1201 | tgggtacattacatgtctgtcgaaggaaggaactgggaactatgctcccttaagttccgc   | 1260 |  |
| QY | 1680 | cccgatgacagaaggtataaaagttcaaatgttgaaacaatggccctcaggtattggagg   | 1739 |  |
| Db | 1261 | cccgatgacagaaggtataaaagttcaaatgttgaaacaatggccctcaggtattggagg   | 1320 |  |
| QY | 1740 | aagtacaaaaaagttccgattatggtacacagatgnaaacctttaagactcgtgaaagt    | 1799 |  |
| Db | 1321 | aagtacaaaaaagttccgattatggtacacagatgnaaacctttaagactcgtgaaagt    | 1380 |  |
| QY | 1800 | attcagcatgtaacaggaatcttgggaaccttacttaatggggaagtgtgagacaactg    | 1859 |  |
| Db | 1381 | attcagcatgtaacaggaatcttgggaaccttacttaatggggaagtgtgagacaactg    | 1440 |  |
| QY | 1860 | ttgattatatttaagaaatcaagcaagcagacataaacatctacccctcaggaatcact    | 1919 |  |
| Db | 1441 | ttgattatatttaagaaatcaagcaagcagacataaacatctacccctcaggaatcact    | 1500 |  |
| QY | 1920 | gattgcctctcttctgtaataaggaagtacccaaaagtgtaaacaatttgaagaatttc    | 1979 |  |
| Db | 1501 | gattgcctctcttctgtaataaggaagtacccaaaagtgtaaacaatttgaagaatttc    | 1560 |  |
| QY | 1980 | ccaatctgcgcaggaagaataatccaataatgaatgacagtgaactgagaagtggcca     | 2039 |  |
| Db | 1561 | ccaatctgcgcaggaagaataatccaataatgaatgacagtgaactgagaagtggcca     | 1620 |  |
| QY | 2040 | actaaatcgaatctctggtccgagccggtatattactctgattctgttaatatggaaga    | 2099 |  |
| Db | 1621 | actaaatcgaatctctggtccgagccggtatattactctgattctgttaatatggaaga    | 1680 |  |
| QY | 2100 | gattcagttccagactcatgttgccctctctcatcctgctacacaagaatctgtgatcaa   | 2159 |  |
| Db | 1681 | gattcagttccagactcatgttgccctctctcatcctgctacacaagaatctgtgatcaa   | 1740 |  |
| QY | 2160 | agaggaacccaataatgtlcaagaagaaggaatgtcatcctgtttctgtatattgatgag   | 2219 |  |
| Db | 1741 | agaggaacccaataatgtlcaagaagaaggaatgtcatcctgtttctgtatattgatgag   | 1800 |  |
| QY | 2220 | aaacgaagctgtgtacacctcaagaagaatatacaacgctttctcccaattcagcttgaggt | 2279 |  |
| Db | 1801 | aaacgaagctgtgtacacctcaagaagaatatacaacgctttctcccaattcagcttgaggt | 1860 |  |
| QY | 2280 | cagcttgagatccagaagtctcaagctcccaacatcatgacagcatlcaaigtgctatgt   | 2339 |  |
| Db | 1861 | cagcttgagatccagaagtctcaagctcccaacatcatgacagcatlcaaigtgctatgt   | 1920 |  |
| QY | 2340 | tttgaatagtttgcagttgtcagttgttgcagtgaagtgtgcatcaactgtatcttaagc   | 2399 |  |
| Db | 1921 | tttgaatagtttgcagttgtcagttgttgcagtgaagtgtgcatcaactgtatcttaagc   | 1980 |  |
| QY | 2400 | attggagcaagagctgaactccttctgctcttctcctcgtgatatccctcaaaccaaa     | 2459 |  |
| Db | 1981 | attggagcaagagctgaactccttctgctcttctcctcgtgatatccctcaaaccaaa     | 2040 |  |
| QY | 2460 | atgtgtcatgaaacacaactacccatattcccatcttcagaggaanaactgtctcatgtcg  | 2519 |  |
| Db | 2041 | atgtgtcatgaaacacaactacccatattcccatcttcagaggaanaactgtctcatgtcg  | 2100 |  |
| QY | 2520 | atggaaaaaccgaagtctatggaattctggggtgccaacactcgaactttggaaacagagc  | 2579 |  |
| Db | 2101 | atggaaaaaccgaagtctatggaattctggggtgccaacactcgaactttggaaacagagc  | 2160 |  |



[illegible][illegible]



Db 4321 ctgactgcctacctcgaattcacccccaaggttggtgcaccagattgcctcagagatg 4480  
 Oy 4800 gaggttcggagctcgagagacagagactcactcagactcagcgagttc 4848  
 |||||  
 Db 4381 gaggttcggagctcgagagacagagactcactcagaggttgccactgc 4429

RESULT 5  
 AAT73164  
 ID AAT73164 standard; CDNA: 9354 BP.  
 AC AAT73164;  
 DT 08-APR-1998 (first entry)  
 XX  
 DE cDNA encoding human B-domain deleted factor VIII.  
 XX  
 KW Post-translational regulatory element; PRE; enhancer II; intronless gene;  
 KW surface antigen gene; cytoplasmic accumulation; targeted delivery;  
 KW near consensus splice sequence; blood coagulation factor; factor VIII;  
 KW factor IX; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 2965..7380  
 FT misc\_feature 5165..5174  
 FT /tag= a  
 FT /tag= b  
 FT /note= "5' near consensus site"  
 FT 5695..5703  
 FT /tag= c  
 FT /note= "5' near consensus site"  
 FT 6320..6328  
 FT /tag= d  
 FT /note= "5' near consensus site"  
 FT 6595..5603  
 FT /tag= e  
 FT /note= "5' near consensus site"  
 FT 7045..7053  
 FT /tag= f  
 FT /note= "5' near consensus site"  
 FT 7143..7152  
 FT /tag= g  
 FT /note= "5' near consensus site"  
 FT 3296..3312  
 FT /tag= h  
 FT /note= "3' near consensus site"  
 FT 4798..4817  
 FT /tag= i  
 FT /note= "3' near consensus site"  
 FT 5023..5045  
 FT /tag= j  
 FT /note= "3' near consensus site"  
 FT 5333..5355  
 FT /tag= k  
 FT /note= "3' near consensus site"  
 FT 5520..5538  
 FT /tag= l  
 FT /note= "3' near consensus site"  
 FT 5604..5632  
 FT /tag= m  
 FT /note= "3' near consensus site"  
 FT 5717..5745  
 FT /tag= n  
 FT /note= "3' near consensus site"  
 FT 6239..6258  
 FT /tag= o  
 FT /note= "3' near consensus site"  
 FT 6658..6682  
 FT /tag= p  
 FT /note= "3' near consensus site"  
 FT 7159..7176  
 FT misc\_feature

FT /tag= q  
 FT /note= "3' near consensus site"  
 FT 7196..7209  
 FT /tag= r  
 FT /note= "3' near consensus site"  
 FT 7289..7315  
 FT /tag= s  
 FT /note= "3' near consensus site"  
 FT 7411..7429  
 FT /tag= t  
 FT /note= "3' near consensus site"  
 FT 7611..8197  
 FT /tag= u  
 FT /note= "PRE sequence"

PN W09733994-A1.  
 XX  
 PD 18-SEP-1997.  
 XX  
 PF 10-MAR-1997: 97WC-US03561.  
 XX  
 PR 11-MAR-1996: 96US-0683839.  
 XX  
 PA (IMMU-) IMMUNE RESPONSE CORP.  
 XX  
 PI Bidingmaier S, Ill CR;  
 XX  
 DR WPI: 1997-470874/43.  
 DR P-PSDB; AAW23414.  
 XX  
 PT Vector for increased expression of intronless genes - comprises  
 PT intronless gene with at least one near consensus splice sequence, a  
 PT promoter and at least one viral cis-acting post-transcriptional  
 PT regulatory element  
 XX  
 PS Example 1: Pages 21-31; 59pp; English.  
 XX  
 CC The present sequence represents human B-domain deleted factor VIII  
 CC cDNA, and a post-translational regulatory element (PRE) of the  
 CC Hepatitis B virus, which is present 3' of the stop codon for factor VIII.  
 CC PRE sequences have been shown to function in cis to increase the  
 CC steady-state levels of surface gene transcripts by facilitating  
 CC cytoplasmic accumulation of these transcripts. The present sequence  
 CC is part of a novel vector, comprising an intronless gene containing  
 CC 1 or more near consensus splice sequences operably linked to a  
 CC promoter sequence so that the gene is transcribed in a cell.  
 CC Intronless gene transcripts which contain near consensus splice site  
 CC sequences are believed to get tied up in the nucleus of the cell where  
 CC splicing occurs, rather than being transported to the cytoplasm where  
 CC they can be translated into proteins. The PRE sequences are transcribed  
 CC along with the gene, causing export of the gene transcript from the  
 CC nucleus into the cytoplasm of the cell. The vector can be used  
 CC to increase the expression of an intronless gene containing at least one  
 CC near consensus splice sites, preferably cDNA encoding a blood coagulation  
 CC factor, particularly factor VIII or IX. The complex allows the targeted  
 CC delivery of the vector to a specific cell, e.g. hepatocytes when the  
 CC ligand is an asialoglycoprotein which binds the asialoglycoprotein  
 CC receptor present on their surface.  
 CC  
 XX  
 SQ Sequence 9354 BP; 2506 A; 2239 C; 2161 G; 2448 T; 0 other;

Query Match 55.5%; Score 4411.6; DB 18; Length 9354;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 4420; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Oy 415 ccacatgcaaatagagctccacacgctctctctctgctgcttgccttgcctta 474  
 |||||  
 Db 2960 ccacatggtttatgagctccacccgctctctctctgcttgccttgcctta 3019  
 |||||  
 Oy 475 gtgcacacgaagaatatactactggttgcaagtgcagtcatatggactatatacaagt 534  
 |||||  
 Db 3020 gtgcacacgaagaatatactactggttgcaagtgcagtcatatggactatatacaagt 3079

QY 535 atctcgtgtagctgctctgtgagcgaagattctctctctagagtgccaaatctttccat 594  
|||||  
Db 3080 atctcgtgtagctgctctgtgagcgaagattctctctctagagtgccaaatctttccat 3139  
QY 595 tcaacaacctcgtctgtacaaaagaactctgtctgtagaattcaacggttccatttca 654  
|||||  
Db 3140 tcaacaacctcgtctgtacaaaagaactctgtctgtagaattcaacggttccatttca 3199  
QY 655 acatcgcgaagccaagcgaaccttgatgggtctgttaggtctctccatccagcgctgag 714  
|||||  
Db 3200 acatcgcgaagccaagcgaaccttgatgggtctgttaggtctctccatccagcgctgag 3259  
QY 715 tttatgtacagctgtgtactaacacttaagaacatgtctcccatctgtcaatcttca 774  
|||||  
Db 3260 tttatgtacagctgtgtactaacacttaagaacatgtctcccatctgtcaatcttca 3319  
QY 775 ctgtctgtgtactctactctggaagaactctctgagagctgtgaatatagtatcagacacagtc 834  
|||||  
Db 3320 ctgtctgtgtactctactctggaagaactctctgagagctgtgaatatagtatcagacacagtc 3379  
QY 835 aaagggaaagaagatgatataaagtcttcctcgtgtggaagccaatgatatgtctgagag 894  
|||||  
Db 3380 aaagggaaagaagatgatataaagtcttcctcgtgtggaagccaatgatatgtctgagag 3439  
QY 895 tctctgaaagagaaatgtgtccaaatgtgaccttgaccactgtgaccttaactatctt 954  
|||||  
Db 3440 tctctgaaagagaaatgtgtccaaatgtgaccttgaccactgtgaccttaactatctt 3499  
QY 955 ctcaatgtgagacctgtgtataaagaacttgaattcagcgctcatgtgagacctactagatgta 1014  
|||||  
Db 3500 ctcaatgtgagacctgtgtataaagaacttgaattcagcgctcatgtgagacctactagatgta 3559  
QY 1015 gagaagggagctctgcccagaagaaagacacagaccttgacaaaattatatacttcttg 1074  
|||||  
Db 3560 gagaagggagctctgcccagaagaaagacacagaccttgacaaaattatatacttcttg 3619  
QY 1075 ctgtatttgatgaaagggaagaagctgtgacatccagaaacaaagaactccttgatgtcagagata 1134  
|||||  
Db 3620 ctgtatttgatgaaagggaagaagctgtgacatccagaaacaaagaactccttgatgtcagagata 3679  
QY 1135 gggatgtcgtacatctgtctggcgctgtgacctaaatgtcacacagtcgaatgttataatata 1194  
|||||  
Db 3680 gggatgtcgtacatctgtctggcgctgtgacctaaatgtcacacagtcgaatgttataatata 3739  
QY 1195 ggtctctgcccagctgtgactgtgacacaggaatacagcttattgtgacatgtgtgaa 1254  
|||||  
Db 3740 ggtctctgcccagctgtgactgtgacacaggaatacagcttattgtgacatgtgtgaa 3799  
QY 1255 tgggacacacatctctgaaatgtgacatcaatattctctcgaaggtcacacatcttctgtgagga 1314  
|||||  
Db 3800 tgggacacacatctctgaaatgtgacatcaatattctctcgaaggtcacacatcttctgtgagga 3859  
QY 1315 accatctgcacggcgctctctgtgaaatctctgcacataacttctcttactgtctcaaaacct 1374  
|||||  
Db 3860 accatctgcacggcgctctctgtgaaatctctgcacataacttctcttactgtctcaaaacct 3919  
QY 1375 tgaatgacctgtgagacagttctactggtttgttcataatctctccacacacatgata 1434  
|||||  
Db 3920 tgaatgacctgtgagacagttctactggtttgttcataatctctccacacacatgata 3979  
QY 1435 tggaaacttatgtcaaaatagacagctgttccagaggaaccccaactacagaaatgaaata 1494  
|||||  
Db 3980 tggaaacttatgtcaaaatagacagctgttccagaggaaccccaactacagaaatgaaata 4039  
QY 1495 atggaagaagcggaagacatgatatgattgtactactgattctgaaatgtatgtgtgcaagt 1554  
|||||  
Db 4040 atggaagaagcggaagacatgatatgattgtactactgattctgaaatgtatgtgtgcaagt 4099  
QY 1555 ttgagtgataaactctctctcttataccaaatctgctcagtttcgacaagaagatctcta 1614  
|||||  
Db 4100 ttgagtgataaactctctctcttataccaaatctgctcagtttcgacaagaagatctcta 4159

QY 1615 aaacttggtatcatcatctgtctgtgaagaggaagacttgggactatgtctcccttagtc 1674  
|||||  
Db 4160 aaacttggtatcatcatctgtctgtgaagaggaagacttgggactatgtctcccttagtc 4219  
QY 1675 tggcccccgtatgacagaagttataaagtccaatatttgaaacaatgtgcccctcagcgagttg 1734  
|||||  
Db 4220 tggcccccgtatgacagaagttataaagtccaatatttgaaacaatgtgcccctcagcgagttg 4279  
QY 1735 gtagaaggtacaaaaaagtccgaatttatgtgtacacagatgaaacctttaagactgtg 1794  
|||||  
Db 4280 gtagaaggtacaaaaaagtccgaatttatgtgtacacagatgaaacctttaagactgtg 4339  
QY 1795 aagctatccaagatgaaatccaagaaatcttgggaaccttacttatgtgggaagttgagaca 1854  
|||||  
Db 4340 aagctatccaagatgaaatccaagaaatcttgggaaccttacttatgtgggaagttgagaca 4399  
QY 1855 caactgtatataatttaagaatccaagcagaacacataaacatctaaccttcacgaa 1914  
|||||  
Db 4400 caactgtatataatttaagaatccaagcagaacacataaacatctaaccttcacgaa 4459  
QY 1915 tcaactggtccgtctcttctgtatataagggagattacccaaaggtgttaaaacatttgaag 1974  
|||||  
Db 4460 tcaactggtccgtctcttctgtatataagggagattacccaaaggtgttaaaacatttgaag 4519  
QY 1975 atttccaattctgtccaggaagaataattcaataataatggacagctgactgtagaagatg 2034  
|||||  
Db 4520 atttccaattctgtccaggaagaataattcaataataatggacagctgactgtagaagatg 4579  
QY 2035 ggcacacataatccaatctctcgtgtgacctgacccgtattacttagtttcgttaatttg 2094  
|||||  
Db 4580 ggcacacataatccaatctctcgtgtgacctgacccgtattacttagtttcgttaatttg 4639  
QY 2095 aagagatcttagcttccaagactcattgtgacctctctctctatctgttacaagaactgtag 2154  
|||||  
Db 4640 aagagatcttagcttccaagactcattgtgacctctctctctatctgttacaagaactgtag 4699  
QY 2155 atcaaaagggaaacacagataatgttcagacaagaaggaatgtcatccgtttctgtatttg 2214  
|||||  
Db 4700 atcaaaagggaaacacagataatgttcagacaagaaggaatgtcatccgtttctgtatttg 4759  
QY 2215 atgagaaccgaagctgtgtacatccacagagaatatacaacggtttctcccaatccagct 2274  
|||||  
Db 4760 atgagaaccgaagctgtgtacatccacagagaatatacaacggtttctcccaatccagct 4819  
QY 2275 gagtgcagcttgagagatccagatgtccaagctccaacataltgacagcaatcaatgtct 2334  
|||||  
Db 4820 gagtgcagcttgagagatccagatgtccaagctccaacataltgacagcaatcaatgtct 4879  
QY 2335 atgtttttgataagtttgtagttgtcagttgtgttgatgtaggtgtgacatctgtatactc 2394  
|||||  
Db 4880 atgtttttgataagtttgtagttgtgtcagttgtgttgatgtaggtgtgacatctgtatactc 4939  
QY 2395 taagaattggagacagactgacttctctgtctctctctctctctctctctctctctctaac 2454  
|||||  
Db 4940 taagaattggagacagactgacttctctgtctctctctctctctctctctctctctctaac 4999  
QY 2455 acaaaatgtctatgaaacacacatcaacctatcccatcttccagtgaaactgtctcta 2514  
|||||  
Db 5000 acaaaatgtctatgaaacacacatcaacctatcccatcttccagtgaaactgtctcta 5059  
QY 2515 tgtcgaatgaaacacacaggtctatgtatctctgtgggtgtgccaaactcagactctgtgaa 2574  
|||||  
Db 5060 tgtcgaatgaaacacacaggtctatgtatctctgtgggtgtgccaaactcagactctgtgaa 5119  
QY 2575 gaggcatacgcgcttacttgaaggtttctgtgttgacaagaacacatgtgtattatag 2634  
|||||  
Db 5120 gaggcatacgcgcttacttgaaggtttctgtgttgacaagaacacatgtgtattatag 5179  
QY 2635 aggaagcttatgaaagataattccaagatctgtcgtgaaatgaaacatgtcatgtgaaacaa 2694  
|||||  
Db 5180 aggaagcttatgaaagataattccaagatctgtcgtgaaatgaaacatgtcatgtgaaacaa 5239  
QY 2695 gaagcttctccagaattccaagacacccctagcactaggtgcaaaagaaatttaattgtcaccc 2754

|||||  
Db 5240 gaagcttccagagattcaagacaccctagactaggcaagaatttaatgccacc 5239  
Oy 2755 caccgctcttgaagcgccatcaacgaggaataactcgtactactcttcagcagatcaag 2814  
Db 5300 caccgctcttgaagcgccatcaacgaggaataactcgtactactcttcagcagatcaag 5359  
Oy 2815 aggaattgactatgatataccatcagcttgaatgaagaagaatttgaatt 2874  
Db 5360 aggaattgactatgatataccatcagcttgaatgaagaagaatttgaatt 5419  
Oy 2875 atgataagatgaataatcaagcccccgagcttccaagaanaaacagacacatttta 2934  
Db 5420 atgataagatgaataatcaagcccccgagcttccaagaanaaacagacacatttta 5479  
Oy 2935 ttgctgcagttgagaagctctgggattatggatgagtagctcccaatgttctaaaga 2994  
Db 5480 ttgctgcagttgagaagctctgggattatggatgagtagctcccaatgttctaaaga 5539  
Oy 2995 acagggctcagaagtggcagttccctcagttccaagaagttgtttccaggaatttactg 3054  
Db 5540 acagggctcagaagtggcagttccctcagttccaagaagttgtttccaggaatttactg 5599  
Oy 3055 atggctccttactcaagcccttataccgttggagaactaaatgaacatttggactcttg 3114  
Db 5600 atggctccttactcaagcccttataccgttggagaactaaatgaacatttggactcttg 5659  
Oy 3115 ggccatataatagaagcagaagttgaagataatatacgttaacttccagaatacagacct 3174  
Db 5660 ggccatataatagaagcagaagttgaagataatatacgttaacttccagaatacagacct 5719  
Oy 3175 ctgctccctactccttctatcttagccttattctttagaagaagatccagagcagaagag 3234  
Db 5720 ctgctccctactccttctatcttagccttattctttagaagaagatccagagcagaagag 5779  
Oy 3235 cagaacctagaaaaaactttgtcaagccataa tgaacccaaacttacttttggaaatgtgc 3294  
Db 5780 cagaacctagaaaaaactttgtcaagccataa tgaacccaaacttacttttggaaatgtgc 5839  
Oy 3295 aacatcatatgtgcacccatataagaatgttgaactgcgaagcttggctattctcttg 3354  
Db 5840 aacatcatatgtgcacccatataagaatgttgaactgcgaagcttggctattctcttg 5899  
Oy 3355 atgttgacctgaaaaaagatgtgcactcagcgcctgatttggcccccttctgtccaca 3414  
Db 5900 atgttgacctgaaaaaagatgtgcactcagcgcctgatttggcccccttctgtccaca 5959  
Oy 3415 ctaaacacatgaacctgtctcatgtggagacaagtgaacagtaagaatttgcctgtttt 3474  
Db 5960 ctaaacacatgaacctgtctcatgtggagacaagtgaacagtaagaatttgcctgtttt 6019  
Oy 3475 tcaacatcttctgtatggagaccaaagctgtgtactccaactgaaataatgaagaacatgtca 3534  
Db 6020 tcaacatcttctgtatggagaccaaagctgtgtactccaactgaaataatgaagaacatgtca 6079  
Oy 3535 gggctccctgcaatatccagatggagaatcccaactttaagaagaaatatacgtcttccatg 3594  
Db 6080 gggctccctgcaatatccagatggagaatcccaactttaagaagaaatatacgtcttccatg 6139  
Oy 3595 caatcaatgtgtacataatgaatatacactactgtgttagtlaatgtgtcagagatcaaga 3654  
Db 6140 caatcaatgtgtacataatgaatatacactactgtgttagtlaatgtgtcagagatcaaga 6199  
Oy 3655 ttctgatgtatctgtctcagatctggcagcaatgaagaataccatctcatattctcagtg 3714  
Db 6200 ttctgatgtatctgtctcagatctggcagcaatgaagaataccatctcatattctcagtg 6259  
Oy 3715 gacatgttctaacgttaagaaaaagaggtataaagtgcacttacaatctctatc 3774  
Db 6260 gacatgttctaacgttaagaaaaagaggtataaagtgcacttacaatctctatc 6319  
Oy 3775 caagtgtttttgagacagttgaaaaatgtataccatccaaagctggaatttggcgggtggaa 3834  
|||||

Db 6320 caagtgtttttgagacagttgaaaaatgtataccatccaaagctggaatttggcgggtggaa 6379  
Oy 3835 gcttatattggcgaacatctatactgtgtggaatgaacacacttttctgtgtacagcaata 3884  
Db 6380 gcttatattggcgaacatctatactgtgtggaatgaacacacttttctgtgtacagcaata 6439  
Oy 3895 agttcagaactcccttggaaatgtctcttgacacatataagatttccagattcaactt 3954  
Db 6440 agttcagaactcccttggaaatgtctcttgacacatataagatttccagattcaactt 6499  
Oy 3955 caagacaataatgaacagttggcccccaagctggccagacttcatattccgataatca 4014  
Db 6500 caagacaataatgaacagttggcccccaagctggccagacttcatattccgataatca 6559  
Oy 4015 atgcttgagacacaaagagcccttcttctgtatcaagtgatctgttgcacccaatga 4074  
Db 6560 atgcttgagacacaaagagcccttcttctgtatcaagtgatctgttgcacccaatga 6619  
Oy 4075 ttattcaagcatcaaacagaggtggccgttcaagaattctccagctctacatctctc 4134  
Db 6620 ttattcaagcatcaaacagaggtggccgttcaagaattctccagctctacatctctc 6679  
Oy 4135 agttatcatcatgtatagttcttgatggagaagttggcagacttatccagaggaattcca 4194  
Db 6680 agttatcatcatgtatagttcttgatggagaagttggcagacttatccagaggaattcca 6739  
Oy 4195 ctggaaaccttaatgtctcttcttggcaatgttgatcttcctggagataaacaacatatt 4254  
Db 6740 ctggaaaccttaatgtctcttcttggcaatgttgatcttcctggagataaacaacatatt 6799  
Oy 4255 ttaacccctcaattatgtctcgatataatccgtttggaccaccaatcatatagattcgca 4314  
Db 6800 ttaacccctcaattatgtctcgatataatccgtttggaccaccaatcatatagattcgca 6859  
Oy 4315 gcaactctgcataagatgtatgtggcgtgtgatttaaatagttgcagcatgtccatttggaa 4374  
Db 6860 gcaactctgcataagatgtatgtggcgtgtgatttaaatagttgcagcatgtccatttggaa 6919  
Oy 4375 tggagaggtaaagcaatatcaatgtcacagatatactgtcttaacttctacttccaatatgt 4434  
Db 6920 tggagaggtaaagcaatatcaatgtcacagatatactgtcttaacttctacttccaatatgt 6979  
Oy 4435 ttggcacactgtctcccttccaanaagctcgacttcaactccaaggagagtaactccttga 4494  
Db 6980 ttggcacactgtctcccttccaanaagctcgacttcaactccaaggagagtaactccttga 7039  
Oy 4495 gacctcaggtgaataatccaaaagatgtgctgcaagtggacttccagagaacaaatgaag 4554  
Db 7040 gacctcaggtgaataatccaaaagatgtgctgcaagtggacttccagagaacaaatgaag 7099  
Oy 4555 tcaacagagtaactacatccagggatataaattctcgtcttaccagcatgtatgtgaagagt 4614  
Db 7100 tcaacagagtaactacatccagggatataaattctcgtcttaccagcatgtatgtgaagagt 7159  
Oy 4615 tccctcatctccagcagatcaagaatgtgccaatgtgactctcttttccaagaatgtgcaag 4674  
Db 7160 tccctcatctccagcagatcaagaatgtgccaatgtgactctcttttccaagaatgtgcaag 7219  
Oy 4675 taaaggttttcaagggaatcaagaactccttcaaacctgtgtgaactctctatagaccac 4734  
Db 7220 taaaggttttcaagggaatcaagaactccttcaaacctgtgtgaactctctatagaccac 7279  
Oy 4735 cgttactgactcgtctactcttccaattcacccccaagatttgggtgcacagatttgccttga 4794  
Db 7280 cgttactgactcgtctactcttccaattcacccccaagatttgggtgcacagatttgccttga 7339  
Oy 4795 ggaatggagttctgtgctgcagagcacagagaccttactgactcgagcaggttc 4848  
Db 7340 ggaatggagttctgtgctgcagagcacagagaccttactgactcgagcaggttc 7393  
|||||

RESULT 6  
AAQ76016

|    |   |  |
|----|---|--|
| ID | AAQ76016 standard: cDNA; 4629 BP.                                     |  |
| XX | AAQ76016;   |  |
| AC |   |  |
| XX | 20-JUL-1995 (first entry)   |  |
| DT |   |  |
| XX | B-domain deleted Factor-VIII.   |  |
| DE |   |  |
| XX | Factor-VIII; blood-clotting; hemophilia A; gene therapy;              |  |
| KW | adenovirus; vector; ss.   |  |
| XX |   |  |
| OS | Homo sapiens.   |  |
| XX |   |  |
| PN | WO9429471-A.  |  |
| XX |   |  |
| PD | 22-DEC-1994.  |  |
| XX |   |  |
| PF | 13-APR-1994; 94WO-US04075.  |  |
| XX |   |  |
| PR | 10-JUN-1993; 93US-0074920.  |  |
| XX | 25-MAR-1994; 94US-0218335.  |  |
| XX |   |  |
| PA | (GENE-) GENETIC THERAPY INC.  |  |
| XX |   |  |
| PI | Connelly S, Kaleko M, Smith T;  |  |
| XX |   |  |
| DR | WPI; 1995-036495/05.  |  |
| DR | P-PSDB; AAR67709.   |  |
| XX |   |  |
| XX | New adenoviral vectors for treatment of haemophilia - contg. a        |  |
| PT | DNA sequence encoding a clotting factor, partic. Factor VIII or       |  |
| PT | Factor IX   |  |
| XX |   |  |
| XX | Disclosure; Fig. 17A-17C; 116pp; English.                             |  |
| XX |   |  |
| CC | Human Factor-VIII cDNA, from which the B domain had been deleted, was |  |
| CC | used to construct recombinant adenovirus vectors that produced        |  |
| CC | therapeutic levels of the clotting factor when administered to an     |  |
| CC | animal host, potentially providing hemophilia A gene therapy.         |  |
| XX |   |  |
| XX |   |  |
| SQ | Sequence 4629 BP; 1318 A; 1051 C; 1018 G; 1242 T; 0 other;            |  |

  

|                       |              |               |                |              |    |
|-----------------------|--------------|---------------|----------------|--------------|----|
| Query Match           | 54.8%;       | Score 4355.4; | DB 16;         | Length 4629; |    |
| Best Local Similarity | 99.0%;       | Pred. No. 0;  |                |              |    |
| Matches 4383;         | Conservative | 0;            | Mismatches 46; | Indels       | 0; |
|                       |              |               |                | Gaps         | 0; |

  

|    |  |     |
|----|--|-----|
| QY | 420 atgcaaatggaactctccacccgtctcttcttggtgaccttttcgatactcgtcttagtgcc   | 479 |
|    |  |     |
| DB | 1 atgcaaatggaactctccacccgtctcttcttggtgaccttttcgatactcgtcttagtgcc     | 60  |
| QY | 480 accaagaatactacctggtgtagtgaactgtaactgtaactgatactgcaagtgatctc      | 539 |
| DB | 61 accaagaatactacctggtgtagtgaactgtaactgtaactgatactgcaagtgatctc       | 120 |
| QY | 540 ggtgaagctgctctgtagaagcaagatttccctcctagatgtagtgcacaacatcttccatc   | 599 |
|    |  |     |
| DB | 121 ggtgaagctgctctgtagaagcaagatttccctcctagatgtagtgcacaacatcttccatc   | 180 |
| QY | 600 acctcaagtcgtgtacaaaagaactcgtttgttagaatcaaggttcaactcttccaacatc    | 659 |
| DB | 181 acctcaagtcgtgtacaaaagaactcgtttgttagaatcaaggttcaactcttccaacatc    | 240 |
| QY | 660 gctaaagcgaagcccccggatggtgctgcgtcaggtccaccatccaggtgagtttat        | 719 |
|    |  |     |
| DB | 241 gctaaagcgaagcccccggatggtgctgcgtcaggtccaccatccaggtgagtttat        | 300 |
| QY | 720 gatacaagtgtcatcacacttaagaacacatggtctccatccctgtcaagtccttatgtgtt   | 779 |
|    |  |     |
| DB | 301 gatacaagtgtgtcatcacacttaagaacacatggtctccatccctgtcaagtccttatgtgtt | 360 |
| QY | 780 ggtgatactacttgaaaagctctctgagggagctggaatgatgatccaagtcagaag        | 839 |
|    |  |     |

|    |      |   |      |
|----|------|---|------|
| Db | 361  | ggctgataccctacgtgaaaagccttcctgaaaggcgctcgatatagatagatacgaacccgtctaaagg  | 420  |
| Qy | 840  | gagaaagaagaatatgataaaagctctccctgtgtgaaagccatacatatagctctgcaaggtctcg     | 899  |
| Db | 421  | gagaaagaagaatgataaaagctcttcctctggtggaagccatacatatgtctgcaaggtctcg        | 480  |
| Qy | 900  | aaagaagaatgtgtccaaatgagcctctgacccacgtgtgcttaactactacatactttctcat        | 959  |
| Db | 481  | aaagaagaatgtgtccaaatgagcctctgacccacgtgtgcttaactactacatactttctcat        | 540  |
| Qy | 960  | gtgagcactggtttaaagaagacttgaattcaacgagcctcatgtgagacccctactagtatgagaga    | 1019 |
| Db | 541  | gtgagcactggtttaaagaagacttgaattcaacgagcctcatgtgagacccctactagtatgagaga    | 600  |
| Qy | 1020 | gggagctctggtccaaaggaaagacacagacctgtgcacaaattatactacttttgcgtgta          | 1079 |
| Db | 601  | gggagctctggtccaaaggaaagacacagacctgtgcacaaattatactacttttgcgtgta          | 660  |
| Qy | 1080 | tttgatgaaagggaaaaagtttggcactcagaaacaaagaactcctcttgatgcaagataagagat      | 1139 |
| Db | 661  | tttgatgaaagggaaaaagtttggcactcagaaacaaagaactcctcttgatgcaagataagagat      | 720  |
| Qy | 1140 | gctgatactgtctcggtcgcttgccctaaatgacacagatcaatgtgttatgtgtaaagagctc        | 1199 |
| Db | 721  | gctgatactgtctcggtcgcttgccctaaatgacacagatcaatgtgttatgtgtaaagagctc        | 780  |
| Qy | 1200 | ctgcagaagctctgcgatttgagatgcacacaggaanaatcagtcgtatattgtagcatgtgatatgtggc | 1259 |
| Db | 781  | ctgcagaagctctgcgatttgagatgcacacaggaanaatcagtcgtatattgtagcatgtgatatgtggc | 840  |
| Qy | 1260 | accactcctgaagatgtcacatcaataatctcctgaaagttcacacatctctgtgtgaggaacct       | 1319 |
| Db | 841  | accactcctgaagatgtcacatcaataatctcctgaaagttcacacatctctgtgtgaggaacct       | 900  |
| Qy | 1320 | cgcaagcggtctcctgggaactctcgccaataacttctcttaacgtgcacaaactcttggt           | 1379 |
| Db | 901  | cgcaagcggtctcctgggaactctcgccaataacttctcttaacgtgcacaaactcttggt           | 960  |
| Qy | 1380 | gacctgtgacagcttctcactcgttttggcattatctctccacacaacatgatagtgcatggaa        | 1439 |
| Db | 961  | gacctgtgacagcttctcactcgttttggcattatctctccacacaacatgatagtgcatggaa        | 1020 |
| Qy | 1440 | gcttatgtcaaaagttagacagctgtctcagaggaaccccaactacgaaatgaaaaataagaa         | 1499 |
| Db | 1021 | gcttatgtcaaaagttagacagctgtctcagaggaaccccaactacgaaatgaaaaataagaa         | 1080 |
| Qy | 1500 | gaagcgggagaagactatgataatgatacttaacgattctcgaatatgatatgtgtgcagtttat       | 1559 |
| Db | 1081 | gaagcgggagaagactatgataatgatacttaacgattctcgaatatgatatgtgtgcagtttat       | 1140 |
| Qy | 1560 | gatgacaactctcctccttatccataatctgcgtcagttgtgccaagagaagatcctaaact          | 1619 |
| Db | 1141 | gatgacaactctcctccttatccataatctgcgtcagttgtgccaagagaagatcctaaact          | 1200 |
| Qy | 1620 | tgggtatacttaccattgtcgtctgaagagagagacttgggactatgctctcccttaagctcgcc       | 1679 |
| Db | 1201 | tgggtatacttaccattgtcgtctgaagagagagacttgggactatgctctcccttaagctcgcc       | 1260 |
| Qy | 1680 | cccgatgacaaagtattataaagtccaattattgaaacaatgagccctcctcagcgggataaggg       | 1739 |
| Db | 1261 | cccgatgacaaagtattataaagtccaattattgaaacaatgagccctcctcagcgggataaggg       | 1320 |
| Qy | 1740 | aagtaacaaaagaagccgattatagcacaagaagatgaaacctttaaagctctgtaagct            | 1799 |
| Db | 1321 | aagtaacaaaagaagccgattatagcacaagaagatgaaacctttaaagctctgtaagct            | 1380 |
| Qy | 1800 | attcaagcatgatacagaagatcttgggaccttcttaactttagtggggaagtgtgaaagacacatg     | 1859 |
| Db | 1381 | attcaagcatgatacagaagatcttgggaccttcttaactttagtggggaagtgtgaaagacacatg     | 1440 |
| Qy | 1860 | ttgatatatttaaagaattcaagaagcagagacaatalaacatctaaccttccacagaaact          | 1919 |
| Db | 1441 | ttgatatatttaaagaattcaagaagcagagacaatalaacatctaaccttccacagaaact          | 1500 |

|    |      |   |      |
|----|------|---|------|
| OY | 1920 | gagtcgcgcgttccttctgtatctcaagagagatccaaaaggtgtgaaaaaatcttgaagatttt     | 1979 |
| Db | 1501 | gatttcgcgcgttccttctgtatctcaagagagatccaaaaggtgtgaaaaaatcttgaagatttt    | 1560 |
| OY | 1980 | ccaatctctgcagagagaataattcaaatataatgtgacgtgtacgtgtgtagagatgtggcca      | 2039 |
| Db | 1561 | ccaatctctgcagagagaataattcaaatataatgtgacgtgtacgtgtgtagagatgtggcca      | 1620 |
| OY | 2040 | actaatcagatccctcgtgtgcctgtaccgcgtataactcagtttcgttaatatgtgagaga        | 2099 |
| Db | 1621 | actaatcagatccctcgtgtgcctgtaccgcgtataactcagtttcgttaatatgtgagaga        | 1680 |
| OY | 2100 | gatttcagcttccagagacatctgtgcctctccatctcgtgtataaagaatctgtatgacca        | 2159 |
| Db | 1661 | gatttcagcttccagagacatctgtgcctctccatctcgtgtataaagaatctgtatgacca        | 1740 |
| OY | 2160 | agaggaacaacagataatgtltaagaacaagagaaatgtcatctcgtttcttcgtatactgtatgag   | 2219 |
| Db | 1741 | agaggaacaacagataatgtltaagaacaagagaaatgtcatctcgtttcttcgtatactgtatgag   | 1800 |
| OY | 2220 | aaccggaagctcgtgtactctcaagagaaataaacaagctttctcccaatccagctcgtgagt       | 2279 |
| Db | 1801 | aaccggaagctcgtgtactctcaagagaaataaacaagctttctcccaatccagctcgtgagt       | 1860 |
| OY | 2280 | cagctttagagatccagagttccagctcccaagctcccaacatcatgtcaagaatcaatgtgtatgt   | 2339 |
| Db | 1861 | cagctttagagatccagagttccagctcccaagctcccaacatcatgtcaagaatcaatgtgtatgt   | 1920 |
| OY | 2340 | tttgatagtttgagttgacgtgttcagtttctgtgtacatgtgagtggtcaatacgtgtacatctagc  | 2399 |
| Db | 1921 | tttgatagtttgagttgacgtgttcagtttctgtgtacatgtgagtggtcaatacgtgtacatctagc  | 1980 |
| OY | 2400 | attggaagacagagactgtactctctctctgtgcctctctctcgtatccatcttaacccaacaaa     | 2459 |
| Db | 1961 | attggaagacagagactgtactctctctctgtgcctctctctcgtatccatcttaacccaacaaa     | 2040 |
| OY | 2460 | atgtgtcctaataagaacacacatccatccatctccatctccatctccagagaacaactgtctcatgtc | 2519 |
| Db | 2041 | atgtgtcctaataagaacacacatccatccatctccatctccatctccagagaacaactgtctcatgtc | 2100 |
| OY | 2520 | attggaaaaacccaagttatagatgttcgtggtgtgcacaactcgaacttcgtgaaaacaagagc     | 2579 |
| Db | 2101 | attggaaaaacccaagttatagatgttcgtggtgtgcacaactcgaacttcgtgaaaacaagagc     | 2160 |
| OY | 2580 | atgacgcgccttactgtgaaggtttctctagtctgtgacaagaacactgtgtatctatcagagac     | 2639 |
| Db | 2161 | atgacgcgccttactgtgaaggtttctctagtctgtgacaagaacactgtgtatctatcagagac     | 2220 |
| OY | 2640 | agttataagaatatcttcaagcatabctgtcgtgtgataaaaacaatgtccatctgtgaaccaagaagc | 2699 |
| Db | 2221 | agttataagaatatcttcaagcatabctgtcgtgtgataaaaacaatgtccatctgtgaaccaagaagc | 2280 |
| OY | 2700 | ttctccccaagattccaagaaccccttagcactatagccaagaagcaatttaatgtcacccccacca   | 2759 |
| Db | 2281 | ttctccccaagattccaagaaccccttagcactatagccaagaagcaatttaatgtcacccccacca   | 2340 |
| OY | 2760 | gtctctgaacaacgcataccaacgaggaataaactcgttactactcttcagttcagatccaagagaa   | 2819 |
| Db | 2341 | gtctctgaacaacgcataccaacgaggaataaactcgttactactcttcagttcagatccaagagaa   | 2400 |
| OY | 2820 | attgacataataatccatatcaatgaattgaatgaagaaggaagattttgaataattatgat        | 2879 |
| Db | 2401 | attgacataataatccatatcaatgaattgaatgaagaaggaagattttgaataattatgat        | 2460 |
| OY | 2880 | gagagatgaanaatcaagagcccccagcttccaagaagaacacgcacatattatattgtc          | 2939 |
| Db | 2461 | gagagatgaanaatcaagagcccccagcttccaagaagaacacgcacatattatattgtc          | 2520 |
| OY | 2940 | gcagttgagagagctctcgtgattatagtgatagtatctcccaatatttctaagaacaag          | 2999 |
| Db | 2521 | gcagttgagagagctctcgtgattatagtgatagtatctcccaatatttctaagaacaag          | 2580 |

|    |      |   |      |
|----|------|---|------|
| QY | 3000 | gccaagatggaatgacgtccctcagctccagaagaatggtttcttcagaagaattacatgagcc  | 3059 |
| Db | 2581 | gtccagatgtggcaatgctccctcagctccagaagaatggtttcttcagaagaattacatgagcc | 2640 |
| QY | 3060 | tcctttactcagcccttatccgttgagaactcaatgaaacctttggagctccctggggcca     | 3119 |
| Db | 2641 | tcctttactcagcccttatccgttgagaactcaatgaaacctttggagctccctggggcca     | 2700 |
| QY | 3120 | ataaagaagagaagatgtgaaataataataatgtaactttcagaatccagccctcgt         | 3179 |
| Db | 2701 | ataaagaagagaagatgtgaaataataataatgtaactttcagaatccagccctcgt         | 2760 |
| QY | 3180 | ccctattccctctatctctagccttatttcttatgagaagatccagggccaagagcagaa      | 3229 |
| Db | 2761 | ccctattccctctatctctagccttatttcttatgagaagatccagggccaagagcagaa      | 2820 |
| QY | 3240 | ccctgaaaaaacctttgcacagccctaatagaaccaaacctattcttgaaaagtccaact      | 3299 |
| Db | 2821 | ccctgaaaaaacctttgcacagccctaatagaaccaaacctattcttgaaaagtccaact      | 2880 |
| QY | 3300 | cataatgaccccaataaagaatgatttgatgcgaagccctgggcctattctcttgatgt       | 3359 |
| Db | 2881 | cataatgaccccaataaagaatgatttgatgcgaagccctgggcctattctcttgatgt       | 2940 |
| QY | 3360 | gaacctggaanaagaatgtgcacccaagccctgatggaacccctctgcttcgccaactaac     | 3419 |
| Db | 2941 | gaacctggaanaagaatgtgcacccaagccctgatggaacccctctgcttcgccaactaac     | 3000 |
| QY | 3420 | aacactgaaacctgtctctatgggagaacaagtgacagatccagaatttgccctgttttcacc   | 3479 |
| Db | 3001 | aacactgaaacctgtctctatgggagaacaagtgacagatccagaatttgccctgttttcacc   | 3060 |
| QY | 3480 | atcctttgatggaaccaaaagcctgttactcaactcagaaataatgaaagaacatcgagggct   | 3539 |
| Db | 3061 | atcctttgatggaaccaaaagcctgttactcaactcagaaataatgaaagaacatcgagggct   | 3120 |
| QY | 3540 | cccttgcaataatccagaatgagaagatcccaactttbaaagagaattatcgcttccatgcaatc | 3599 |
| Db | 3121 | cccttgcaataatccagaatgagaagatcccaactttbaaagagaattatcgcttccatgcaatc | 3180 |
| QY | 3600 | aattgctacaataatgagatacaactcctgtcttagtaatgcttcagaatccaagaatttga    | 3659 |
| Db | 3181 | aattgctacaataatgagatacaactcctgtcttagtaatgcttcagaatccaagaatttga    | 3240 |
| QY | 3660 | tgtatctctgcacagcagggcagcaacaatgaaacatccattccattccattccagtgagcat   | 3719 |
| Db | 3241 | tgtatctctgcacagcagggcagcaacaatgaaacatccattccattccattccagtgagcat   | 3300 |
| QY | 3720 | gtgttcaactgtacgaaaaaagaagsgataaaatgycacgtgtacacatctcatccaggt      | 3779 |
| Db | 3301 | gtgttcaactgtacgaaaaaagaagsgataaaatgycacgtgtacacatctcatccaggt      | 3360 |
| QY | 3780 | gttttttgagaagatgggaatggttaccatcccaagcctggaatttggcgggttggaatgctt   | 3839 |
| Db | 3361 | gttttttgagaagatgggaatggttaccatcccaagcctggaatttggcgggttggaatgctt   | 3420 |
| QY | 3840 | attggcagacatctacaatgtctgggaatgagacaacttttctgtgttacaagcaataatggt   | 3899 |
| Db | 3421 | attggcagacatctacaatgtctgggaatgagacaacttttctgtgttacaagcaataatggt   | 3480 |
| QY | 3900 | cagagctccctctgggaatgagctttcttgagacacatagaagatttcagattacagcttcagaa | 3959 |
| Db | 3481 | cagagctccctctgggaatgagctttcttgagacacatagaagatttcagattacagcttcagaa | 3540 |
| QY | 3960 | caataatggaacagtgagccccaagaagctggccacagatctcatattccggatcaataatgcc  | 4019 |
| Db | 3541 | caataatggaacagtgagccccaagaagctggccacagatctcatattccggatcaataatgcc  | 3600 |
| QY | 4020 | tgagagcaacaagagccctttctcttgagacaagatgagatctgttgagccaatgatattt     | 4079 |
| Db | 3601 | tgagagcaacaagagccctttctcttgagacaagatgagatctgttgagccaatgatattt     | 3660 |
| QY | 4080 | cacggagatcaagaaccccggtgcccgttcagaagatttccagcgccttaatatcttcagattt  | 4139 |

```
Db      3661 caccgcaatcagaccgagggctgcctcgaagatcttccagcctctacatctcaagttc 3720
OY      4140 atcatctgatactctctgatacgtgaagaagatgagcagacttctcagagaatctccactga 4199
Db      3721 atcaatctgatactctctgatacgtgaagaagatgagcagacttctcagagaatctccactga 3780
OY      4200 accttaatgctctctcttctgcaatctgatacctctggaataaacaacataatttttaac 4259
Db      3781 accttaatgctctctcttctgcaatctgatacctctggaataaacaacataatttttaac 3840
OY      4260 cctccaattattgctcgtatatacatccgtttgaccccaactcttatagatctgcagcact 4319
Db      3841 cctccaattattgctcgtatatacatccgtttgaccccaactcttatagatctgcagcact 3900
OY      4320 ctccgcatgagatgatacgtgagcttgatattaatagttgcacatgacctttggaaatgag 4379
Db      3901 ctccgcatgagatgatacgtgagcttgatattaatagttgcacatgacctttggaaatgag 3960
OY      4380 agtaaaacataatcagatctgacagatctctcctcactccttaccataatgtttgccc 4439
Db      3961 agtaaaacataatcagatctgacagatctctcctcactccttaccataatgtttgccc 4020
OY      4440 acctgtgtcctccttcaaaaagctcgcagcttccactccaagggagatgaatgctctgagacct 4499
Db      4021 acctgtgtcctccttcaaaaagctcgcagcttccactccaagggagatgaatgctctgagacct 4080
OY      4500 caggtgataatccaaaagatgagctgcgaagtgagactccgaagacataatgaatccaca 4559
Db      4081 caggtgataatccaaaagatgagctgcgaagtgagactccgaagacataatgaatccaca 4140
OY      4560 gggagtaactactcagaggagtaaaaatctctgcttaccagacatgatagtgaaggagttcctc 4619
Db      4141 gggagtaactactcagaggagtaaaaatctctgcttaccagacatgatagtgaaggagttcctc 4200
OY      4620 atctccagcagtcagatgagtcacatgagatctcttttttcagaatgagcaagtaag 4679
Db      4201 atctccagcagtcagatgagtcacatgagatctcttttttcagaatgagcaagtaag 4260
OY      4680 gttttcagaaggaatacgaagactccttcaacactgtgtgtaactcctcgaacccaactga 4739
Db      4261 gttttcagaaggaatacgaagactccttcaacactgtgtgtaactcctcgaacccaactga 4320
OY      4740 ctgactcgtactactcgaatccaccccaagatgtgggtgacacagatgtccctcgaagatg 4799
Db      4321 ctgactcgtactactcgaatccaccccaagatgtgggtgacacagatgtccctcgaagatg 4380
OY      4800 gaaggtctggtctgctgagagcagaagactcctactgacctcgaagcgaatgc 4848
Db      4381 gaaggtctggtctgctgagagcagaagactcctactgagggtggtccactgc 4429

RESULT 7
AAV23339
ID      AAV23339 standard; DNA: 4670 BP.
XX
AC      AAV23339:
XX
DT      17-AUG-1998 (first entry)
XX
DE      Human Factor-VIII gene lacking central B domain.
XX
KW      Factor-VIII; blood clotting; human; synthetic gene; codon usage; ss.
OS      Homo sapiens.
XX
PN      WO9812207-A1.
XX
PD      26-MAR-1998.
XX
PF      18-SEP-1997; 97WO-US16639.
XX
```

```
PR      20-SEP-1996; 96US-0717294.
PA      (GEHO ) GEN HOSPITAL CORP.
XX
PI      Haas J, Seed B:
XX
DR      WPI: 1998-217200/19.
XX
PT      New synthetic eukaryotic gene(s) - in which non-preferred or less
PT      preferred codon(s) are replaced to provide high level expression in
PT      mammalian cell(s)
XX
PS      Example 3; Fig 12; 92pp; English.
XX
CC      This gene codes for a human Factor-VIII protein that lacks the
CC      central B domain (amino acids 760-1639) of the native protein. In
CC      a novel, claimed synthetic gene (see AAV23288), non-preferred or
CC      less preferred codons of the native gene are replaced by codons
CC      favored by highly expressed human genes to provide high-level
CC      expression in mammalian cells. The synthetic gene was assembled
CC      from 29 pairs of oligonucleotides (see AAV2340-97) which served as
CC      PCR templates. Synthetic genes of the invention (see also
CC      AAV23289-91) are used for production of recombinant proteins in
CC      mammalian cells at levels of at least 500% of those obtained using
CC      the natural genes. They can also be used in gene therapy.
XX
SQ      Sequence 4670 BP; 1325 A; 1072 C; 1036 G; 1237 T; 0 other;

Query Match      54.8%; Score 4351.2; DB 19; Length 4670;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 4359; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

OY      477 gccaccagaataactactcgtgtgcagtggaactgtcatgagactatagtcaaatgtgat 536
Db      84 gccaccagaataactactcgtgtgcagtggaactgtcatgagactatagtcaaatgtgat 143
OY      537 ctgcgtgagctgtcctgtgagacgaagatcttcctcctagatgccaatacttttccatc 566
Db      144 ctgcgtgagctgtcctgtgagacgaagatcttcctcctagatgccaatacttttccatc 203
OY      557 aacacctgagctgcgtgatacaaaaagactctgtttgtaaatcaacgtttccctttccaa 656
Db      204 aacacctgagctgcgtgatacaaaaagactctgtttgtaaatcaacgtttccctttccaa 263
OY      657 atcgttaagccaagccaacctctgagatgtgtcgttaagttcctaccatccagctgagtt 716
Db      264 atcgttaagccaagccaacctctgagatgtgtcgttaagttcctaccatccagctgagtt 323
OY      717 tatgatacagtggtcaatatacaacttaagaacatggtcttccatccctgacgttcatgct 776
Db      324 tatgatacagtggtcaatatacaacttaagaacatggtcttccatccctgacgttcatgct 383
OY      777 gttgtgtatctctactcgtgaagagcttctcgaaggagctgaatataatgatacagacagtc 836
Db      384 gttgtgtatctctactcgtgaagagcttctcgaaggagctgaatataatgatacagacagtc 443
OY      837 agggagaagaagatgataaagcttctccctgtgtggaagccaatacatatgtctggcaggtc 896
Db      444 agggagaagaagatgataaagcttctccctgtgtggaagccaatacatatgtctggcaggtc 503
OY      897 ctgaaagagaatggttccaatgtgctctgacccactgtgcttacctatcatatcttctc 956
Db      504 ctgaaagagaatggttccaatgtgctctgacccactgtgcttacctatcatatcttctc 563
OY      957 catgtgagcctgtaaaaaagacttaattcaaggctctattggagccctactagratatgaga 1016
Db      564 catgtgagcctgtaaaaaagacttaattcaaggctctattggagccctactagratatgaga 623
OY      1017 gaaggagctgtgccaaggaagaacacagacacttgacaatttatactacttctgtc 1076
Db      624 gaaggagctgtgccaaggaagaacacagacacttgacaatttatactacttctgtc 683
```

QY 1077 gtaattgataaggagaaagtctggcaactcagaacaaagaactccttgatcagatag 1136  
|||||  
Db 684 gtattgatgaaggagaaagtctggcaactcagaacaaagaactccttgatcagatag 743  
QY 1137 gatctgcattctgctggcctggcctaaatgcaacagtcattgtttgttaacagg 1196  
|||||  
Db 744 gatctgcattctgctggcctggcctaaatgcaacagtcattgtttgttaacagg 803  
QY 1197 tctctgcagatctgattgattgattgcacaggaatcattcttgatggcctgattggatg 1256  
|||||  
Db 804 tctctgcagatctgattgattgattgcacaggaatcattcttgatggcctgattggatg 863  
QY 1257 ggcacacctcctgaagtcgacccctcaatactcctgaaggtcacacattctctgtgagAAC 1316  
|||||  
Db 864 ggcacacctcctgaagtcgacccctcaatactcctgaaggtcacacattctctgtgagAAC 923  
QY 1317 catgcagagcgtctctgtgaaatctgcaccaatacttcttactctgtcacaacctctg 1376  
|||||  
Db 924 catgcagagcgtctctgtgaaatctgcaccaatacttcttactctgtcacaacctctg 983  
QY 1377 atggactctggacagttcttactggtttgttcatactcttccacacacatgattgcatg 1436  
|||||  
Db 984 atggactctggacagttcttactggtttgttcatactcttccacacacatgattgcatg 1043  
QY 1437 gaagcttatgtlcaaaagtacagcgtgtccagaggaaaccccaactacgaatgaaanaaat 1496  
|||||  
Db 1044 gaagcttatgtlcaaaagtacagcgtgtccagaggaaaccccaactacgaatgaaanaaat 1103  
QY 1497 gaagaagcggaaagactatgattgattcttactgattctgtaaatgtgtgttcoagttt 1556  
|||||  
Db 1104 gaagaagcggaaagactatgattgattcttactgattctgtaaatgtgtgttcoagttt 1163  
QY 1557 gatgatcacacactctctcttcttattccaatcgctcagtttcgaagaagccctccaa 1616  
|||||  
Db 1164 gatgatcacacactctctcttcttattccaatcgctcagtttcgaagaagccctccaa 1223  
QY 1617 acttggtatcattacatctgtctgtcgtgaagaggagactgggactatgtctcccttactctc 1676  
|||||  
Db 1224 acttggtatcattacatctgtctgtcgtgaagaggagactgggactatgtctcccttactctc 1283  
QY 1677 gcccccgtacaggaagtctataaagtccaattttgaaacatgtgcctcagcggaattgtgt 1736  
|||||  
Db 1284 gcccccgtacaggaagtctataaagtccaattttgaaacatgtgcctcagcggaattgtgt 1343  
QY 1737 aggaagatcacaaaagtcgcgaattatggtacatacacagatgaacactttaagactcgtgaa 1796  
|||||  
Db 1344 aggaagatcacaaaagtcgcgaattatggtacatacacagatgaacactttaagaccgtgaa 1403  
QY 1797 gctattcagatgaatlcaggaatccttggaacttacttatagtgggaagtctgagacaca 1856  
|||||  
Db 1404 gctattcagatgaatlcaggaatccttggaacttacttatagtgggaagtctgagacaca 1463  
QY 1857 ctgtttgattatatttaagaatcgaagcagacacataacatctacccctcaggaatc 1916  
|||||  
Db 1464 ctgtttgattatatttaagaatcgaagcagacacataacatctacccctcaggaatc 1523  
QY 1917 acgtgatctccgtccttctgtatcttaagagagattacaaaaggtgtlaaaacattggaagat 1976  
|||||  
Db 1524 acgtgatctccgtccttctgtatcttaagagagattacaaaaggtgtlaaaacattggaagat 1583  
QY 1977 ttccaattctgcagaggaataatataaataaataatgacagtgagctgtagaagatg 2036  
|||||  
Db 1584 ttccaattctgcagaggaataatataaataaataatgacagtgagctgtagaagatg 1443  
QY 2037 ccaactaatacagatcctcggtgcctgcgaaccgcttatctactgattctgttaataatgag 2096  
|||||  
Db 1644 ccaactaatacagatcctcggtgcctgcgaaccgcttatctactgattctgttaataatgag 1703  
QY 2097 agagatctaatcgaagaccatttgccctctcctacatctgcgtacaagaagattgtagat 2156  
|||||  
Db 1704 agagatctaatcgaagaccatttgccctctcctacatctgcgtacaagaagattgtagat 1763  
QY 2157 caaagagaaacagataatgttcagacaagaagattgtcatctgttttctgtattgtat 2216

Db 1764 caaagagaaacagataatgttcagacaagaagattgtcatctgttttctgtattgtat 1823  
|||||  
QY 2217 gaagaaccgaagctgtgtacccacagagataatacaagcttctccccaactcagcttga 2276  
|||||  
Db 1824 gaagaaccgaagctgtgtacccacagagataatacaagcttctccccaactcagcttga 1883  
QY 2277 gtgcagcttgagatcccaaggtcccaagctcccaactcaatgacagatcaatgtgctat 2336  
|||||  
Db 1884 gtgcagcttgagatcccaaggtcccaagctcccaactcaatgacagatcaatgtgctat 1943  
QY 2337 gtctttgatagttctgcagttgtcagttgtgttcattgagattgacatgtgtatctta 2396  
|||||  
Db 1944 gtctttgatagttctgcagttgtcagttgtgttcattgagattgacatgtgtatctta 2003  
QY 2397 agcatvgagacagactgtactccttctctgtcttctctctctctctctctctctctct 2456  
|||||  
Db 2004 agcatvgagacagactgtactccttctctgtcttctctctctctctctctctctctct 2063  
QY 2457 aaaaigtctataagaacacacactcaacttcccatctcagagagaactgtctcatg 2516  
|||||  
Db 2064 aaaaigtctataagaacacacactcaacttcccatctcagagagaactgtctcatg 2123  
QY 2517 tctgatgaaaaaccaggtctcatgtatctgtgggtgcacacactcagacacttctcgaaaca 2576  
|||||  
Db 2124 tctgatgaaaaaccaggtctcatgtatctgtgggtgcacacactcagacacttctcgaaaca 2183  
QY 2577 ggcatacgcgcttactgaaagtttctagtgtgtacagaagaacgtgtgattctcag 2636  
|||||  
Db 2184 ggcatacgcgcttactgaaagtttctagtgtgtacagaagaacgtgtgattctcag 2243  
QY 2637 gacagttataagaatattttacgatactgtctgagttaaaaaaatccattgaaaccaag 2696  
|||||  
Db 2244 gacagttataagaatattttacgatactgtctgagttaaaaaaatccattgaaaccaag 2303  
QY 2697 agcttctcccaagaattccaagaaccccttagacataagccaagaatlttaatgtcacccca 2756  
|||||  
Db 2304 agcttctcccaagaattccaagaaccccttagacataagccaagaatlttaatgtcacccca 2263  
QY 2757 ccagcttctgaaagccatccaagcggaataactcgttactactcttcaagtacataag 2816  
|||||  
Db 2364 ccagcttctgaaagccatccaagcggaataactcgttactactcttcaagtacataag 2423  
QY 2817 gaattgactatgatatcacatatacagttgaaatgaagaagaattttgacattat 2876  
|||||  
Db 2424 gaattgactatgatatcacatatacagttgaaatgaagaagaattttgacattat 2483  
QY 2877 gatgagatgaataatcagaagcccccagcgttccaagaagaacacagacacttttatt 2936  
|||||  
Db 2484 gatgagatgaataatcagaagcccccagcgttccaagaagaacacagacacttttatt 2543  
QY 2937 gctgcagatgagaagctcctgtggaattatggaatgagttagctcccaatgttctaagaac 2996  
|||||  
Db 2544 gctgcagatgagaagctcctgtggaattatggaatgagttagctcccaatgttctaagaac 2603  
QY 2997 agggctcagaatgtgcagtttccctcagttccaagaagaagttgttttcagaagaatttactgt 3056  
|||||  
Db 2604 agggctcagaatgtgcagtttccctcagttccaagaagaagttgttttcagaagaatttactgt 2663  
QY 3057 ggtctcttactcagccctbatacgttggagaactlaaataatgttggaactctgtggg 3116  
|||||  
Db 2664 ggtctcttactcagccctbatacgttggagaactlaaataatgttggaactctgtggg 2723  
QY 3117 ccaatataagacagaagtgtgaagataatataatgttacttccagaatacagcgctct 3176  
|||||  
Db 2724 ccaatataagacagaagtgtgaagataatataatgttacttccagaatacagcgctct 2783  
QY 3177 cgtccctattccttctacttactacactattcttctgaagaagatcagagcagaagaca 3236  
|||||  
Db 2784 cgtccctattccttctacttactacactattcttctgaagaagatcagagcagaagaca 2843  
QY 3237 gaaactagaaaaaacttgtlcaaagcctlaatgaaaccaaaacttacttcttgaaagtgc 3296  
|||||



Db 2844 gaaacctgaaaaaaacttctgcaagccctaaatgaacccaaacttacttcttgaaagtgcga 2903  
QY 3297 catcatactgacccacacataaagatgagtttctgactgcaagccctgggtctattctctgat 3356  
Db 2904 catcatactgacccacacataaagatgagtttctgactgcaagccctgggtctattctctgat 2963  
QY 3357 gttgaacctgaaaaaagtctgactcaagccctgatttggaacccctctgtctgtccacact 3416  
Db 2964 gttgaacctgaaaaaagtctgactcaagccctgatttggaacccctctgtctgtccacact 3023  
QY 3417 aaacacatgaaacctgtctgcatgaggagacaagtgaacagfacaggaatttgctgtcttttc 3476  
Db 3024 aaacacatgaaacctgtctgcatgaggagacaagtgaacagfacaggaatttgctgtcttttc 3083  
QY 3477 accatcttgatgagaccaaagaactgtgtactctacatgaaaaatagaaaaagaagaactgcag 3536  
Db 3084 accatcttgatgagaccaaagaactgtgtactctacatgaaaaatagaaaaagaagaactgcag 3143  
QY 3537 gctccctgcgaataatccagatggaagatcccaactttaagaagaattatcgcttccatgca 3596  
Db 3144 gctccctgcgaataatccagatggaagatcccaactttaagaagaattatcgcttccatgca 3203  
QY 3597 atcaatgcatcaataatggaatcacactactggttagtaatggtctcagagatccaagaatt 3656  
Db 3204 atcaatgcatcaataatggaatcacactactggttagtaatggtctcagagatccaagaatt 3263  
QY 3657 cgaatgtatctgtctcagcatgtggcagaatgaaaaacatccatctcatcttcacgttga 3716  
Db 3264 cgaatgtatctgtctcagcatgtggcagaatgaaaaacatccatctcatcttcacgttga 3323  
QY 3717 catgtgttcaactgttaacgaaaaaagaaggagataaaatgtgcactgtacaatctctacca 3776  
Db 3324 catgtgttcaactgttaacgaaaaaagaaggagataaaatgtgcactgtacaatctctacca 3383  
QY 3777 ggtgttttttagagacagtggaaattttaccatccaaagctggaatttggcgggttggaatgc 3836  
Db 3384 ggtgttttttagagacagtggaaattttaccatccaaagctggaatttggcgggttggaatgc 3443  
QY 3837 cttaattggcagacatctactacatgtcgtggaatgagacacacttttctgtgttaacagcataag 3896  
Db 3444 cttaattggcagacatctactacatgtcgtggaatgagacacacttttctgtgttaacagcataag 3503  
QY 3897 tgtcagagctccctcgtggaatgtgcttctggaacacattagagaatttcagattacagcttca 3956  
Db 3504 tgtcagagctccctcgtggaatgtgcttctggaacacattagagaatttcagattacagcttca 3563  
QY 3957 ggaacaatgagacagtgtggccccaagaactggcagactcttaattccggatccaatcat 4016  
Db 3564 ggaacaatgagacagtgtggccccaagaactggcagactcttaattccggatccaatcat 3623  
QY 4017 gacctgagacccaagagcccttctcttggaatcaaggttggaatctgtgttgacccaatgatt 4076  
Db 3624 gacctgagacccaagagcccttctcttggaatcaaggttggaatctgtgttgacccaatgatt 3683  
QY 4077 attcacagcatccaagaccagggtgtcccggtcagaagtcttccagccttcaatctcttag 4136  
Db 3684 attcacagcatccaagaccagggtgtcccggtcagaagtcttccagccttcaatctcttag 3743  
QY 4137 ttatactcatgtatagttcttgatggaagaagtgtgcagagatttcgaggaattccact 4196  
Db 3744 ttatactcatgtatagttcttgatggaagaagtgtgcagagatttcgaggaattccact 3803  
QY 4197 ggaaccttaatgttctctcttgcaatgttgatcatcatctggaataaacaacataatttt 4256  
Db 3804 ggaaccttaatgttctctcttgcaatgttgatcatcatctggaataaacaacataatttt 3863  
QY 4257 aaacctccaattatgtctgatatacatccgtttgcacccaactcatatagcatctgcagc 4316  
Db 3864 aaacctccaattatgtctgatatacatccgtttgcacccaactcatatagcatctgcagc 3923  
QY 4317 acctctgcagatgagttgagtggtgtgatttaaatagttgcaacatgcatgttggaattg 4376  
Db 3924 acctctgcagatgagttgagtggtgtgatttaaatagttgcaacatgcatgttggaattg 3983

QY 4377 gagagtaagaacatatcatagatgcacaagattactgtctcatcctactcttacaatatgttt 4436  
Db 3984 gagagtaagaacatatcatagatgcacaagattactgtctcatcctactcttacaatatgttt 4043  
QY 4437 gccacctgtctctctcccaaaagctcgacttcaactccaaggagagtaatgctctgaga 4496  
Db 4044 gccacctgtctctctcccaaaagctcgacttcaactccaaggagagtaatgctctgaga 4103  
QY 4497 cctcaggtggaataatccaaaagatggtgtcaagtgtgacttccagagaagaatgaaagtgc 4556  
Db 4104 cctcaggtggaataatccaaaagatggtgtcaagtgtgacttccagagaagaatgaaagtgc 4163  
QY 4557 acaaggagtaactactcaaggagtaaaatctctgtcttaaccagatgtatgtgaaggagtgc 4616  
Db 4164 acaaggagtaactactcaaggagtaaaatctctgtcttaaccagatgtatgtgaaggagtgc 4223  
QY 4617 ctcatctccagcagtcacagatggtgccatcagtgagactctcttttctcgaatgtgcaagtta 4676  
Db 4224 ctcatctccagcagtcacagatggtgccatcagtgagactctcttttctcgaatgtgcaagtta 4283  
QY 4677 aaggttttccagggaataatccaagactccttcaacacctgtgttgaaactctcagaaccacg 4736  
Db 4284 aaggttttccagggaataatccaagactccttcaacacctgtgttgaaactctcagaaccacg 4343  
QY 4737 ttactgactcgctacactcgaattcaacccccaagatgtgggtgcacagattgacctgag 4796  
Db 4344 ttactgactcgctacactcgaattcaacccccaagatgtgggtgcacagattgacctgag 4403  
QY 4797 atggaagttctgtggtcgtgcagagcagacgactctcactgactgcagcagatc 4848  
Db 4404 atggaagttctgtggtcgtgcagagcagacgactctcactgaggtgtgccaatgc 4455

RESULT 8  
AAV19581  
ID AAV19581 standard; cDNA: 4832 BP.  
XX  
XX AAV19581;  
AC  
XX  
XX 06-AUG-1998 (first entry)  
DT  
XX  
XX Human factor VIII beta-domain deleted SQN deletion cDNA sequence.  
DE  
XX  
XX Replication defective; recombinant retrovirus; RRV; therapeutic protein;  
KW haemophilia; thrombosis; hypercoagulable disorder; liver disease; human;  
KW hepatitis; thalassemia; phenylketonuria; Lesch-Nyhan syndrome; diabetes;  
KW cystic fibrosis; Duchenne's Muscular Dystrophy; hypercholesterolemia;  
KW hypopituitarism; adenine deaminase deficiency; HIV infection; anaemia;  
KW Gaucher's syndrome; high blood pressure; Alzheimer's disease, autoimmune;  
KW inflammatory disease; factor VIII; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX W09800541-A2.  
FN  
XX  
XX 08-JAN-1998.  
PD  
XX  
XX 02-JUL-1997; 97WO-US11784.  
PE  
XX  
XX 04-JUN-1997; 97US-0869309.  
PR  
XX  
XX 03-JUL-1996; 96US-0645601.  
PR  
XX  
XX 13-AUG-1996; 96US-0696381.  
PA  
XX  
XX (CHIR ) CHIRON CORP.  
XX  
XX Allen JR, Barber JR, Boder M, Chang SMW, Chong K;  
PI De LA VEGA D, Depoloni J, Greengard J, Hsu DC, Ibanez CE;  
PI Jolly DJ, Lee R, Mittelestaedt DM, Prusak CE, Respass JG;  
XX  
XX WPI: 1998-086966/08.  
DR  
XX  
XX P-PSDB: AAW46246.

PT New replication defective recombinant retro-viruses - which can be  
PT administered to provide long term systemic expression of therapeutic  
PT protein in blood, useful in, e.g. treating hyper-coagulable  
PT disorders

XX Example 26; Pages 210-213; 272pp; English.

XX This cDNA encodes the beta-domain deleted SON deletion protein of human  
CC factor VIII. This is used in the construction of recombinant retroviral  
CC vectors expressing human factor VIII. The invention provides the  
CC preparation of replication defective recombinant retrovirus (RV)  
CC expressing a therapeutic protein. The RV preparation is resistant to  
CC degradation by human complement and is capable of inducing long term  
CC systemic expression of the therapeutic protein when administered  
CC intravenously to a human. The long term systemic expression results in a  
CC measurable level of the therapeutic protein being produced in the blood  
CC of the human for a period of at least 30 days after the administration of  
CC the RV vector preparation. RV's can be used for in vivo delivery of  
CC therapeutic protein to treat, e.g. haemophilia A, haemophilia B,  
CC thrombosis, hypercoagulable disorders, liver diseases such as hepatitis,  
CC disorders such as thalassemia, phenylketonuria, Lesch-Nyhan syndrome,  
CC severe combined immunodeficiency (SCID), cystic fibrosis, Duchenne's  
CC Muscular Dystrophy, inherited emphysema, familial hypercholesterolemia,  
CC diabetes, hypopituitarism, adenine deaminase deficiency, alpha1-  
CC antitrypsin deficiency, Gaucher's syndrome, anaemia, infections such as  
CC HIV infection, high blood pressure, Alzheimer's disease, autoimmune or  
CC inflammatory disease or graft versus host disease. RV's are capable of  
CC surviving inactivation in human serum thereby allowing efficient gene  
CC transfer over prolonged periods of time.

XX Sequence 4832 BP; 1381 A; 1089 C; 1055 G; 1307 T; 0 other;

Query Match 54.4%; Score 4321.6; DB 19; Length 4832;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 4379; Conservative 0; Mismatches 9; Indels 42; Gaps 1;

QY 419 catccaatagagctccacccctgctctctctctctgctcttgatctgcttattg 478  
DB 71 catgcaatagagctccacccctgctctctctctctgctcttgatctgcttattg 130  
QY 479 caccagaagatactacctggtgagtgagctgcatgcatgagctatataatgagtgatct 538  
DB 131 caccagaagatactacctggtgagtgagctgcatgcatgagctatataatgagtgatct 190  
QY 539 cggtagactgctgtagagcaaatcttctctctagagtgccaaatctttccattcaa 598  
DB 191 cggtagactgctgtagagcaaatcttctctctagagtgccaaatctttccattcaa 250  
QY 599 cactcagctgtagcaaaaagactctgtttagaattcaaggttcaactttcaacat 658  
DB 251 cactcagctgtagcaaaaagactctgtttagaattcaaggttcaactttcaacat 310  
QY 659 cgttaagcacaagccacccctgagatggtctgtagtccctaccatccaggctgaggtta 718  
DB 311 cgttaagcacaagccacccctgagatggtctgtagtccctaccatccaggctgaggtta 370  
QY 719 tgatacagtggttattcaacttaagaacaatggtctccatccctgtagcttctatgctgt 778  
DB 371 tgatacagtggttattcaacttaagaacaatggtctccatccctgtagcttctatgctgt 430  
QY 779 tgggttctcctactggaagctctctgagggagctggaatgatgatataagccagttcaag 838  
DB 431 tgggttctcctactggaagctctctgagggagctggaatgatgatataagccagttcaag 490  
QY 839 ggaagaagaatgataaagtctctccctggttggaagcacaataatgctcggcaggtccct 898  
DB 491 ggaagaagaatgataaagtctctccctggttggaagcacaataatgctcggcaggtccct 550  
QY 899 gaaagaaatggttcaatggtcctctgacccaatggtgcttcaactatcatctttctca 958  
DB 551 gaaagaaatggttcaatggtcctctgacccaatggtgcttcaactatcatctttctca 610

QY 959 tctgacactggtataaagacttgaaatcaagcctcatggtgagccctactagatgtaga 1018  
DB 611 tctgacactggtataaagacttgaaatcaagcctcatggtgagccctactagatgtaga 670  
QY 1019 agggagctctggccaaggaagaaagacacagaccttgccacaatatatactacttttgcgt 1078  
DB 671 agggagctctggccaaggaagaaagacacagaccttgccacaatatatactacttttgcgt 730  
QY 1079 atttgatgaaggaagaagtgtgacatcagaacaaagaactcttgatgacgataagga 1138  
DB 731 atttgatgaaggaagaagtgtgacatcagaacaaagaactcttgatgacgataagga 790  
QY 1139 tgcgcacctctgctggccttgctgcttaaatgcaacagctcaatggttatgaaacaggtc 1198  
DB 791 tgcgcacctctgctggccttgctgcttaaatgcaacagctcaatggttatgaaacaggtc 850  
QY 1199 tctgcaagcttgatgtgagtcacacaggaatacagctcatgagatgagtgagtgag 1258  
DB 851 tctgcaagcttgatgtgagtcacacaggaatacagctcatgagatgagtgagtgagtgag 910  
QY 1259 caccactcctgaaatgacataatctcctcgaagtgcaacacattctgtgagaacca 1318  
DB 911 caccactcctgaaatgacataatctcctcgaagtgcaacacattctgtgagaacca 970  
QY 1319 tgcgcagagctgctcttggaatactcgcgaataacttctctactgcttaaacacttgat 1378  
DB 971 tgcgcagagctgctcttggaatactcgcgaataacttctctactgcttaaacacttgat 1030  
QY 1379 ggaacttggacagtttctactgttttgcataatcttccacacaatatgataggcagta 1438  
DB 1031 ggaacttggacagtttctactgttttgcataatcttccacacaatatgataggcagta 1090  
QY 1439 agcttatgtcaaatgatacagctgtccagaggaaccccaactacgaatgataaataatga 1498  
DB 1091 agcttatgtcaaatgatacagctgtccagaggaaccccaactacgaatgataaataatga 1150  
QY 1499 agaagcggaaagactatgatactgatactactgatacttgcgaatgagtgagtgagtgag 1558  
DB 1151 agaagcggaaagactatgatactgatactactgatacttgcgaatgagtgagtgagtgag 1210  
QY 1559 tgatgaacactctcctctcttcaatccaaatcgctcagctgtgccaagaagcatctcaaac 1618  
DB 1211 tgatgaacactctcctctcttcaatccaaatcgctcagctgtgccaagaagcatctcaaac 1270  
QY 1619 ttggtatcatatacttctgctcgtgagaagggagactggtgactatgctccctagctctgc 1678  
DB 1271 ttggtatcatatacttctgctcgtgagaagggagactggtgactatgctccctagctctgc 1330  
QY 1679 ccccgatgacagaaggtataaagaatcaaatcttgaaacatggtccctcagcgagttgtag 1738  
DB 1331 ccccgatgacagaaggtataaagaatcaaatcttgaaacatggtccctcagcgagttgtag 1390  
QY 1739 gaagtacaaaagaatccgattatgcatatcacacagatgaaacctttaagactctgaaagc 1798  
DB 1391 gaagtacaaaagaatccgattatgcatatcacacagatgaaacctttaagactctgaaagc 1450  
QY 1799 tatccagcagatgaaatcagaatcttgggaacttactcttaatggggagagtttgagacacact 1858  
DB 1451 tatccagcagatgaaatcagaatcttgggaacttactcttaatggggagagtttgagacacact 1510  
QY 1859 gttgattatatttaagaatacagaacagacacataataactacaccccaaggaatcac 1918  
DB 1511 gttgattatatttaagaatacagaacagacacataataactacaccccaaggaatcac 1570  
QY 1919 tgatgctgctcttctgattacaaggaatgataccaaaggtgtaaaaactttgaagattt 1978  
DB 1571 tgatgctgctcttctgattacaaggaatgataccaaaggtgtaaaaactttgaagattt 1630  
QY 1979 tccaattctgcccaggaagaataatcaataataatgagcagtgagtgagaaagtggcc 2038  
DB 1631 tccaattctgcccaggaagaataatcaataataatgagcagtgagtgagaaagtggcc 1690  
QY 2039 aactaaatcagatcctcgtggtccgacccgctattactagttctgtaatatgagagag 2098

Db 1691 aactaaatcaagatcctcgttgcttgaaccgcatctactagtttcgttaataatgagag 1750  
QY 2099 agactagcttcaggaacccatcttgccctctctactctctgtaacaaagaaatcgtatgata 2158  
Db 1751 agactagcttcaggaacccatcttgccctctctactctctgtaacaaagaaatcgtatgata 1810  
QY 2159 aagaggaacaccagataatgttcagacaagaagatgttcaaccgtttctctgtatacttgatga 2218  
Db 1811 aagaggaacaccagataatgttcagacaagaagatgttcaaccgtttctctgtatacttgatga 1870  
QY 2219 gaaccgaagctgtgtaacctcacagagaataatacaagctttctcccaatccagctcgagat 2278  
Db 1871 gaaccgaagctgtgtaacctcacagagaataatacaagctttctcccaatccagctcgagat 1930  
QY 2279 gcaactggaagatccagaggttccaagctccaacatcatgccaagcatcaatggtatgt 2338  
Db 1931 gcaactggaagatccagaggttccaagctccaacatcatgccaagcatcaatggtatgt 1990  
QY 2339 ttctgatagtttgcaagttgtcagtttgttttgcaatgaggttggaatactgttatacttctaa 2398  
Db 1991 ttctgatagtttgcaagttgtcagtttgttttgcaaggttggaatactgttatacttctaa 2050  
QY 2399 catctggagcaagaactgtaacct 2458  
Db 2051 catctggagcaagaactgtaacct 2110  
QY 2459 aatggtctatgaagaacaactccaactctccatctccaagagaagaactgttcatgtc 2518  
Db 2111 aatggtctatgaagaacaactccaactctccatctccaagagaagaactgttcatgtc 2170  
QY 2519 gatvgaaaaaccaggtctatggaatctvgggttgcacacaactcaagaactcttcggaacagag 2578  
Db 2171 gatvgaaaaaccaggtctatggaatctvgggttgcacacaactcaagaactcttcggaacagag 2230  
QY 2579 catgaccgctctactggaagtttctagtttgcagacaagaacactgttgattatacgaagga 2638  
Db 2231 catgaccgctctactggaagtttctagtttgcagacaagaacactgttgattatacgaagga 2290  
QY 2639 cagttatgaagaataatctcagatacttgcgtgaataaacaatgcatctggaacccaagag 2698  
Db 2291 cagttatgaagaataatctcagatacttgcgtgaataaacaatgcatctggaacccaagag 2350  
QY 2699 ctctccccaagaatccaagacacccctagcaactagccaagaacaaatlaabgccaccacc 2758  
Db 2351 ctctccccaaga-----cccaacc 2368  
QY 2759 agtcttgaacgcacatcaacgggaataactcgtactactcttcaagtcagatcaagaagga 2818  
Db 2369 agtcttgaacgcacatcaacgggaataactcgtactactcttcaagtcagatcaagaagga 2428  
QY 2819 aatgtaataatgataacatatcaagttgaaatggaagaaggaatttgcacattatga 2878  
Db 2429 aatgtaataatgataacatatcaagttgaaatggaagaaggaatttgcacattatga 2488  
QY 2879 tgaagatgaaataatcaagcccccgacgtcttcaaaagaaaacgaacatattatctgc 2938  
Db 2489 tgaagatgaaataatcaagcccccgacgtcttcaaaagaaaacgaacatattatctgc 2548  
QY 2939 tgaagtgaagaagctctgggaattatggaatgagtccccaactggttctaagaagaag 2998  
Db 2549 tgaagtgaagaagctctgggaattatggaatgagtccccaactggttctaagaagaag 2608  
QY 2999 ggtcacaagtgcaggtccctcagttccaagaaggttcttccaaggaattactgacgtg 3058  
Db 2609 ggtcacaagtgcaggtccctcagttccaagaaggttcttccaaggaattactgacgtg 2668  
QY 3059 ctctcttactcagccctataacgttggagaactaaatgaacaatttgggaactcctctgggccc 3118  
Db 2669 ctctcttactcagccctataacgttggagaactaaatgaacaatttgggaactcctctgggccc 2728  
QY 3119 atataataagcgagaagtgtgaagataataatcatggttaacttcagaataacagcctctcg 3178

Db 2729 atataataagcgagaagtgtgaagataataatcatggttaacttcagaataacagcctctcg 2788  
QY 3179 tcccatctccctctacttctcagcctattctctatbaagaaatcagagccaaggaaga 3238  
Db 2789 tcccatctccctctacttctcagcctattctctatbaagaaatcagagccaaggaaga 2848  
QY 3239 accatgaaaaaacttcttgaagcctaataatgaacccaacttacttltggaagtgcaca 3298  
Db 2849 accatgaaaaaacttcttgaagcctaataatgaacccaacttacttltggaagtgcaca 2908  
QY 3299 tcatatggaaccacttaagaatgagtttgactgcgaaccccggtctatttctctgagt 3358  
Db 2909 tcatatggaaccacttaagaatgagtttgactgcgaaccccggtctatttctctgagt 2968  
QY 3359 tgaacttgaaaaagaatgagcaactcagcgtgaattggaacctctctgctgcacactaa 3418  
Db 2969 tgaacttgaaaaagaatgagcaactcagcgtgaattggaacctctctgctgcacactaa 3028  
QY 3419 cacaactgaacctgctcatatggaagaacaagtgaagtaacaggaattgtctgttttcaac 3478  
Db 3029 cacaactgaacctgctcatatggaagaacaagtgaagtaacaggaattgtctgttttcaac 3088  
QY 3479 catctttgatgagaccacaagaagctgtgtacttcaactgaaataatggaagaagaatcgaagcc 3538  
Db 3089 catctttgatgagaccacaagaagctgtgtacttcaactgaaataatggaagaagaatcgaagcc 3148  
QY 3539 tccctgcaataatccagatggaagatcccaactttaaagaagatlatcgtctccatgcaat 3598  
Db 3149 tccctgcaataatccagatggaagatcccaactttaaagaagatlatcgtctccatgcaat 3208  
QY 3599 caatgtgtaacataatgataacatactacctggtctagtaatggtcgaagatcaagatctcg 3658  
Db 3209 caatgtgtaacataatgataacatactacctggtctagtaatggtcgaagatcaagatctcg 3268  
QY 3659 atggtatctgctcagacatggtggagcaaatgaaacaatccatctctatccattcaatgga 3718  
Db 3269 atggtatctgctcagacatggtggagcaaatgaaacaatccatctctatccattcaatgga 3328  
QY 3719 tgtgttcaactgtacgaaaaaaagaggaatataaaatgagcaactgtacaaactctatccag 3778  
Db 3329 tgtgttcaactgtacgaaaaaaagaggaatataaaatgagcaactgtacaaactctatccag 3388  
QY 3779 tgtttttgaagacagtggaaatgttaccatccaagctvgaaattvgcggttggaaatgctc 3838  
Db 3389 tgtttttgaagacagtggaaatgttaccatccaagctvgaaattvgcggttggaaatgctc 3448  
QY 3839 tatitggagaagatctacatctgcttgggaatgagacacatttcttctgttgaacgaataagt 3898  
Db 3449 tatitggagaagatctacatctgcttgggaatgagacacatttcttctgttgaacgaataagt 3508  
QY 3899 tcaagactccctctgggaatgagcttctggaacacattagagatttcaagatttaacagctcag 3958  
Db 3509 tcaagactccctctgggaatgagcttctggaacacattagagatttcaagatttaacagctcag 3568  
QY 3959 acaataatggaacagtggcccccaagcttggccaagacttcaatcccggaatcaatcgaatgc 4018  
Db 3569 acaataatggaacagtggcccccaagcttggccaagacttcaatcccggaatcaatcgaatgc 3628  
QY 4019 ctggaagcaaccaagagcccttcttcttgaataaggttgaatctgttggcaaccaatgattat 4078  
Db 3629 ctggaagcaaccaagagcccttcttcttgaataaggttgaatctgttggcaaccaatgattat 3668  
QY 4079 tcaaggaatcaagaaccagaggtgcctcgtaagaagtltcccgagctctcaactctcagtt 4138  
Db 3689 tcaaggaatcaagaaccagaggtgcctcgtaagaagtltcccgagctctcaactctcagtt 3748  
QY 4139 tatcatcatatagttcttgaatggaagaagtggcagaacttatacgaaggaatctccactggt 4198  
Db 3749 tatcatcatatagttcttgaatggaagaagtggcagaacttatacgaaggaatctccactggt 3808  
QY 4199 aaccttaatggtctctcttggcaagtgtgattcaatcttgggaataaacaacataatttttaa 4258  
Db 3809 aaccttaatggtctctcttggcaagtgtgattcaatcttgggaataaacaacataatttttaa 3868

QY 4259 cccccaattatgctgatalacacgcgttgcaccaactcattatagatcgcagcac 4318  
|||||  
Db 3869 cccccaattatgctgatalacacgcgttgcaccaactcattatagatcgcagcac 3928  
4319 tcttcgcatgaggttgatggcgtgtgatttaaatagttgcagcatgaggaatgga 4378  
|||||  
Db 3929 tcttcgcatgaggttgatggcgtgtgatttaaatagttgcagcatgaggaatgga 3988  
4379 gagtaagcaatatacagatgcacagatctactgctcactctactcctaataatgttgc 4438  
|||||  
Db 3989 gagtaagcaatatacagatgcacagatctactgctcactctactcctaataatgttgc 4048  
4439 cactgtgctccctccaagaactgcactcaccctccaaggaggaatgctcctgagacc 4498  
|||||  
Db 4049 cactgtgctccctccaagaactgcactcaccctccaaggaggaatgctcctgagacc 4108  
4499 tcaggtgataataccaaaagatggtgcagtggaacttccagaagaacatgaaatgcac 4558  
|||||  
Db 4109 tcaggtgataataccaaaagatggtgcagtggaacttccagaagaacatgaaatgcac 4168  
4559 agggatgatactcagaggaataaattctgtctaccagatgatagtgaaggatcct 4618  
|||||  
Db 4169 agggatgatactcagaggaataaattctgtctaccagatgatagtgaaggatcct 4228  
4619 catctccagcagtcagaatgagcactcagtgactctcttttcagagatgcaagtaaa 4678  
|||||  
Db 4229 catctccagcagtcagaatgagcactcagtgactctcttttcagagatgcaagtaaa 4288  
4679 ggttttcagggaaatcaagaactcctcacaacbtgtgtgaactctcagaccacgtt 4738  
|||||  
Db 4289 ggttttcagggaaatcaagaactcctcacaacbtgtgtgaactctcagaccacgtt 4348  
4739 actgactcgtactccttgaattacccccagagttggtgtcacagattgcctgaagat 4798  
|||||  
Db 4349 actgactcgtactccttgaattacccccagagttggtgtcacagattgcctgaagat 4408  
4799 ggaagttctggtcgtgcagacagacactctactcgaagcgaattc 4848  
|||||  
Db 4409 ggaagttctggtcgtgcagacagacactctactcgaagcgaattc 4458

RESULT 9  
AAV15338 standard; DNA; 4832 BP.  
ID AAV15338:  
AC AAV15338:  
XX  
XX 20-JUL-1998 (first entry)  
DE Human Factor VIII SQN deletion mutant DNA.  
KW Factor VIII; blood clotting; haemophilia A; gene therapy;  
KW retrovirus; vector; human; ss.  
XX  
XX Homo sapiens.  
OS Synthetic.  
OS  
FH Key Location/Qualifiers  
FT CDS 72..445  
FT /tag= a  
XX  
XX MO9800542-AZ.  
XX  
XX 08-JAN-1998.  
XX  
XX 02-JUL-1997; 97MO-US11785.  
XX  
XX 04-JUN-1997; 97US-0869309.  
XX 03-JUN-1996; 96US-0645601.  
XX 13-AUG-1996; 96US-0696381.  
XX  
XX (CHIR ) CHIRON CORP.

XX  
PI Allen JR, Barber JR, Boder M, Chang SM, Chong K;  
PI De la Vega D, Depolo NJ, Greengard J, Hsu DC, Ibanez CE;  
PI Jolly DJ, Mittlestaedt DM, Prussak CE, Respass JG;  
DR  
DR WPI; 1998-086967/08.  
DR P-SDB: AAM44372.  
XX  
PT New replication defective recombinant retroviruses - which express B  
PT domain-deleted human factor VIII or human factor IX for the  
PT treatment of haemophilia  
XX  
PS Claim 6; Page 174-175; 236pp; English.  
XX  
CC This DNA sequence includes a coding region for the B domain  
CC deletion mutant SQN (see AAM44372) of human factor VIII. The SQN  
CC mutant is created by fusing Ser-743 to Gln-1638 of native factor  
CC VIII (see AAM44373) to form a Ser-Gln-Asn (SQN) link between the A2  
CC and A3 factor VIII domains. When compared to plasmid-derived  
CC factor VIII, the SQN deletion does not influence the in vivo  
CC pharmacokinetics, but the reduced size of the molecule appears to  
CC decrease proteolytic degradation. The invention relates to  
CC preparations of replication defective recombinant retrovirus (RV)  
CC expressing a B domain-deleted human factor VIII protein, where the  
CC recombinant RV is capable of infecting human cells, is resistant  
CC to degradation by human complement and is capable of inducing  
CC long-term (at least 30 days and up to 6 months or longer  
CC post-injection) systemic expression of factor VIII when  
CC administered to a haemophilia A patient.  
XX  
SQ Sequence 4832 BP; 1381 A; 1089 C; 1055 G; 1307 T; 0 other;

Query Match 54.4%; Score 4321.6; DB 19; Length 4832;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 4379; Conservative 0; Mismatches 9; Indels 42; Gaps 1;

QY 419 catccaatgagagctccacactcctctctctctgtgcttcttgatctgattagtc 478  
|||||  
Db 71 catccaatgagagctccacactcctctctctctgtgcttcttgatctgattagtc 130

QY 479 caccagaagatactacctggtgagtgagactgtcatggaactatgtcaaatgtact 538  
|||||  
Db 131 caccagaagatactacctggtgagtgagactgtcatggaactatgtcaaatgtact 190

QY 539 cgtgtgactcctgtgtgagcgaagattcctcctagagtgccaaatctttccatcaa 598  
|||||  
Db 191 cgtgtgactcctgtgtgagcgaagattcctcctagagtgccaaatctttccatcaa 250

QY 599 cactcagtcgtgtgacaaaagactcgtttgtagaattcaagttacactttccaacat 658  
|||||  
Db 251 cactcagtcgtgtgacaaaagactcgtttgtagaattcaagttacactttccaacat 310

QY 659 cgttaagcacaagccaccctgagatggtctgtcgtagtcctaccctcaggctgagttta 718  
|||||  
Db 311 cgttaagcacaagccaccctgagatggtctgtcgtagtcctaccctcaggctgagttta 370

QY 719 tgatacagtggtcaattacacttaagaacatggtcctccactcgtcagttcatgctgt 778  
|||||  
Db 371 tgatacagtggtcaattacacttaagaacatggtcctccactcgtcagttcatgctgt 430

QY 779 tgggtatcctactcgtgaaagcttctgagggagctggaatgatgatgaagccaatgaag 838  
|||||  
Db 431 tgggtatcctactcgtgaaagcttctgagggagctggaatgatgatgaagccaatgaag 490

QY 839 ggaagaagaatgataaagtcttcctcgtgtggaagacatacatatgtcggcaggtcct 898  
|||||  
Db 491 ggaagaagaatgataaagtcttcctcgtgtggaagacatacatatgtcggcaggtcct 550

QY 899 gaaagaagaatggtccaaatggtccttgaccacactgtgcttaactacatacttttca 958  
|||||  
Db 551 gaaagaagaatggtccaaatggtccttgaccacactgtgcttaactacatacttttca 610

|    |      |  |      |
|----|------|--|------|
| QY | 959  | tqvgagccvtgttaaagacttgaattcagagccctacattvgagccctactagatgtataga     | 1018 |
| Db | 611  | tgtvgagccvtgttaaagacttgaattcagagccctacattvgagccctactagatgtataga    | 670  |
| QY | 1019 | aggagatctvgccaaggaagaaagacacagacccttgcacaatttatactacttttcgct       | 1078 |
| Db | 671  | aggagatctvgccaaggaagaaagacacagacccttgcacaatttatactactttttgctgt     | 730  |
| QY | 1079 | atttgaatgaagggaaaaagcttgcacatcagaaacaaaagactccttgaatgcaggaataga    | 1138 |
| Db | 731  | atttgaatgaagggaaaaagcttgcacatcagaaacaaaagactccttgaatgcaggaataga    | 790  |
| QY | 1139 | tgtgtgcatactgcctcvggcccvggacctaaatgacacacgtcaatgtgttatgaaacagctc   | 1198 |
| Db | 791  | tgtgtgcatactgcctcvggcccvggacctaaatgacacacgtcaatgtgttatgaaacagctc   | 850  |
| QY | 1199 | tctgcacaggtctgatgtgagtcacaggaatactagcttatgtgagatgtgatgtgaatggg     | 1258 |
| Db | 851  | tctgcacaggtctgatgtgagtcacaggaatactagcttatgtgagatgtgatgtgaatggg     | 910  |
| QY | 1259 | caaccactccgaaaggtgcacatccaatactccttcgaagtcacacatcttctgtgaggaacca   | 1318 |
| Db | 911  | caaccactccgaaaggtgcacatccaatactccttcgaagtcacacatcttctgtgaggaacca   | 970  |
| QY | 1319 | tctgcagagcgctccttggaaatctcgcaccaatacttccctactgtcctccaacacctctgat   | 1378 |
| Db | 971  | tctgcagagcgctccttggaaatctcgcaccaatacttccctactgtcctccaacacctctgat   | 1030 |
| QY | 1379 | ggagccttggagcagtttctactgttttgcatactcttccacccaataatgtgtgcatgga      | 1438 |
| Db | 1031 | ggagccttggagcagtttctactgttttgcatactcttccacccaataatgtgtgcatgga      | 1090 |
| QY | 1439 | agactatgtcaaaagtacagacagctgtccacagaggaaaccccaactacgaatgtgaaataatga | 1498 |
| Db | 1091 | agactatgtcaaaagtacagacagctgtccacagaggaaaccccaactacgaatgtgaaataatga | 1150 |
| QY | 1499 | agaagcggaaagactatgatagtatgacttctactgatacttcgaaatgtgatagtgtcaggttga | 1558 |
| Db | 1151 | agaagcggaaagactatgatagtatgacttctactgatacttcgaaatgtgatagtgtcaggttga | 1210 |
| QY | 1559 | tgtatgacaactctccctcctttatccaattcgttcaagtttgcagaagagacccctaaac      | 1618 |
| Db | 1211 | tgtatgacaactctccctcctttatccaattcgttcaagtttgcagaagagacccctaaac      | 1270 |
| QY | 1619 | tltggtacatctacatgtctgtctgaagaggagacatgtgagactatgctccctatgacctgcg   | 1678 |
| Db | 1271 | tltggtacatctacatgtctgtctgaagaggagacatgtgagactatgctccctatgacctgcg   | 1330 |
| QY | 1679 | ccccgatgacagaagttataaaagtccaatatltgacaatvgcccctcagcggatgtgtag      | 1738 |
| Db | 1331 | ccccgatgacagaagttataaaagtccaatatltgacaatvgcccctcagcggatgtgtag      | 1390 |
| QY | 1739 | gaagtatacaaaaagttccgatttatgtgatatcacagatgaaacctttagactcgttgaagc    | 1798 |
| Db | 1391 | gaagtatacaaaaagttccgatttatgtgatatcacagatgaaacctttagactcgttgaagc    | 1450 |
| QY | 1799 | tactcagcatgaaatcagaaatctctggagaccttcttatctggggaaagtgtgagacacact    | 1858 |
| Db | 1451 | tactcagcatgaaatcagaaatctctggagaccttcttatctggggaaagtgtgagacacact    | 1510 |
| QY | 1859 | gttgatataatttaagatatcaagacaagacaagacataatactacacccctcaggaatcac     | 1918 |
| Db | 1511 | gttgatataatttaagatatcaagacaagacaagacataatactacacccctcaggaatcac     | 1570 |
| QY | 1919 | tgtatgtcgtctccttgttatccaagagagattaccaaagaagtgttaaacatttgaagattc    | 1978 |
| Db | 1571 | tgtatgtcgtctccttgttatccaagagagattaccaaagaagtgttaaacatttgaagattc    | 1630 |
| QY | 1979 | tccaattctgcagggagaaatatctcaaatbtaaaatggacagctgtgataagaatctggccc    | 2038 |
| Db | 1631 | tccaattctgcagggagaaatatctcaaatbtaaaatggacagctgtgataagaatctggccc    | 1690 |
| QY | 2039 | aactaaatcagatcctctcgtgcgacctgacccgctatctactcagtttcgttataatgagag    | 2098 |

|    |      |  |      |
|----|------|--|------|
| Db | 1691 | aactaaacagatccctccggtgctcgagccgcgtatctccctagtttcgtttaataatgagag    | 1750 |
| Qy | 2099 | aaatactaaccttcagaactcattctggccctccctcccatctctcaagaagaaactcgtatgaca | 2158 |
| Db | 1751 | agatctagcttcagagactcattggtccctccctcccatctgtaaaagaatctgtatgaca      | 1810 |
| Qy | 2155 | aagaggaacacagataatgtctcagaacaagagaaatgtcatccctgtttctgtattgtatga    | 2218 |
| Db | 1811 | aagaggaacacagataatgtctcagaacaagagaaatgtcatccctgtttctgtattgtatga    | 1870 |
| Qy | 2219 | gaaccgaagctcgtgtacccctcaagaagaaatatacaagccttccccaatccagctgaggt     | 2278 |
| Db | 1871 | gaaccgaagctcgtgtacccctcaagaagaaatatacaagccttccccaatccagctgaggt     | 1930 |
| Qy | 2279 | gaagctttagagatcccaagattcccaagcctccaacatcatgacaacagatcaatgctatggt   | 2338 |
| Db | 1931 | gaagctttagagatcccaagattcccaagcctccaacatcatgacaacagatcaatgctatggt   | 1990 |
| Qy | 2339 | ttttagatagtttgcagttgtctcagttgttctgaca tgaaggtgacatactggtatattctaa  | 2398 |
| Db | 1991 | ttttagatagtttgcagttgtctcagttgttctgaca tgaaggtgacatactggtatattctaa  | 2050 |
| Qy | 2399 | catttgagacacagactgacttccctctctgtcctctctctctctgtatatacttcaacacaa    | 2458 |
| Db | 2051 | catttgagacacagactgacttccctctctgtcctctctctctctgtatatacttcaacacaa    | 2110 |
| Qy | 2459 | aattgtctatgaagaacacactccacatcccatctcccatctcaagagaaactgctttcatgtc   | 2518 |
| Db | 2111 | aattgtctatgaagaacacactccacatcccatctcccatctcaagagaaactgctttcatgtc   | 2170 |
| Qy | 2519 | gatggaaaaacccaaggtctatggaattctggggtgccaacaactagaactctcggaacaagag   | 2578 |
| Db | 2171 | gatggaaaaacccaaggtctatggaattctggggtgccaacaactagaactctcggaacaagag   | 2230 |
| Qy | 2579 | catgaacgccttacttgaagtttctctagttgtgccaagaacactggtgtattattacagaga    | 2638 |
| Db | 2231 | catgaacgccttacttgaagtttctctagttgtgccaagaacactggtgtattattacagaga    | 2290 |
| Qy | 2639 | caagtatgaagaatatttccaacatactgtgcgaatgaanaacaaatgcatatgtaaccaagaag  | 2698 |
| Db | 2291 | caagtatgaagaatatttccaacatactgtgcgaatgaanaacaaatgcatatgtaaccaagaag  | 2350 |
| Qy | 2699 | cttctcccaagaattcaagaacacccctagcacatgagcaaaaagcaatttaatgccccaccc    | 2758 |
| Db | 2351 | cttctcccaaga-----ccccaccc-----ccccaccc                             | 2368 |
| Qy | 2759 | agctcttgaacgcctcaacaaagggaagaataactgctactactctcagtcagatcaagaagga   | 2818 |
| Db | 2369 | agctcttgaacgcctcaacaaagggaagaataactgctactactctcagtcagatcaagaagga   | 2428 |
| Qy | 2819 | aattgacatataatataccatatactgaattgaaatgaagaaggaagatttgaacattatga     | 2878 |
| Db | 2429 | aattgacatataatataccatatactgaattgaaatgaagaaggaagatttgaacattatga     | 2488 |
| Qy | 2879 | tgaagatgaanaatacagaagcccccgcagcttccaaaagaanaacacagacataattatgtc    | 2938 |
| Db | 2489 | tgaagatgaanaatacagaagcccccgcagcttccaaaagaanaacacagacataattatgtc    | 2548 |
| Qy | 2939 | tgcagttgagaagcctcgtggattaatgagatgagtagctccccaatggttctaagaacag      | 2998 |
| Db | 2549 | tgcagttgagaagcctcgtggattaatgagatgagtagctccccaatggttctaagaacag      | 2608 |
| Qy | 2999 | ggctcagaaggtgacaggtctccctcaagttccaagaagaagtgttttccaagaaattatctgag  | 3058 |
| Db | 2609 | ggctcagaaggtgacaggtctccctcaagttccaagaagaagtgttttccaagaaattatctgag  | 2668 |
| Qy | 3059 | ctcccttactaagcctctataacgttgaagaacaaatgaaacttgggaacctccggyggcc      | 3118 |
| Db | 2669 | ctcccttactaagcctctataacgttgaagaacaaatgaaacttgggaacctccggyggcc      | 2728 |
| Qy | 3119 | atataaagcagaagttgaagaataatcatatgcttaacttccaagaactcagcctctcg        | 3178 |

Dh 2729 atalataagacagaagltgagataalaatcatgttaacttccagaaatcagccctctcg 2788  
QY 3179 tccattatctctattcttaagccttatttcttaaggaagaatcagagcgcaagacaga 3238  
Dh 2789 tccattatctctattcttaagccttatttcttaaggaagaatcagagcgcaagacaga 2848  
QY 3239 acccagaaaacttctgcaagcctaataagcaaaaacttacttctggaagtgcaaca 3298  
Dh 2849 acccagaaaacttctgcaagcctaataagcaaaaacttacttctggaagtgcaaca 2908  
QY 3299 tcataatgaccccaataagaatgagttgactgcaagcctggcctatttctctgagt 3358  
Dh 2909 tcataatgaccccaataagaatgagttgactgcaagcctggcctatttctctgagt 2968  
QY 3359 tgacctggaaaaaagtctgacacagcctgagtggacccttctgtctgcccactaa 3418  
Dh 2969 tgacctggaaaaaagtctgacacagcctgagtggacccttctgtctgcccactaa 3028  
QY 3419 cacactgaacctctcatatggagacagaagtgacagtaacaggaattgctctgttttcaac 3478  
Dh 3029 cacactgaacctctcatatggagacagaagtgacagtaacaggaattgctctgttttcaac 3088  
QY 3479 catctttgagagaccaaagctgttacttcaactgaaaataatggaaagaaactgcaagggc 3538  
Dh 3089 catctttgagagaccaaagctgttacttcaactgaaaataatggaaagaaactgcaagggc 3148  
QY 3539 tccctgtaatatccagatlgagaatgccactttaaagagaattatgtcttccatgcaat 3598  
Dh 3149 tccctgtaatatccagatlgagaatgccactttaaagagaattatgtcttccatgcaat 3208  
QY 3599 caatgctacataatlgataacacactgctgttagtaatgctcaggaatcagaagatctcg 3658  
Dh 3209 caatgctacataatlgataacacactgctgttagtaatgctcaggaatcagaagatctcg 3268  
QY 3659 atggatctgtcagacatggtggcagcaatlgaaaacatccatcttaattcaattcagtgagca 3718  
Dh 3269 atggatctgtcagacatggtggcagcaatlgaaaacatccatcttaattcaattcagtgagca 3328  
QY 3719 tgtgttaactgttagaaaaaagagagtaataatgacctgtgacacatctcttccagc 3778  
Dh 3229 tgtgttaactgttagaaaaaagagagtaataatgacctgtgacacatctcttccagc 3388  
QY 3779 tgtttttgagcaagltggaatgtttacatccaaagctggaatttggcggtgtggaatgct 3838  
Dh 3389 tgtttttgagcaagltggaatgtttacatccaaagctggaatttggcggtgtggaatgct 3448  
QY 3839 taattggcgagcatctacatgctggtgagtaagcaacttcttctgtgttacaagcaataagtg 3898  
Dh 3449 taattggcgagcatctacatgctggtgagtaagcaacttcttctgtgttacaagcaataagtg 3508  
QY 3899 tcaagctcccttgggaatgtgctcttgacacatagaagatttcaagatttcaagttcagctcagc 3958  
Dh 3509 tcaagctcccttgggaatgtgctcttgacacatagaagatttcaagatttcaagttcagc 3568  
QY 3959 acaatatgagcaagtggtggcccaaaagctggtgcaagatcatatctcggaatcaaatcagtc 4018  
Dh 3569 acaatatgagcaagtggtggcccaaaagctggtgcaagatcatatctcggaatcaaatcagtc 3628  
QY 4019 ctggagacacaagagcccttcttcttgatcaagtgagatctgttggcaccaatgatlat 4078  
Dh 3629 ctggagacacaagagcccttcttcttgatcaagtgagatctgttggcaccaatgatlat 3688  
QY 4079 tcaaggaataaagcccaaggtgtgcccgttcagaagtttccagcctctacatctctcagtt 4138  
Dh 3689 tcaaggaataaagcccaaggtgtgcccgttcagaagtttccagcctctacatctctcagtt 3748  
QY 4139 tatcatcatgtatagttctgagtggaagaagtggtgcaagactatctgagaagaattccacttg 4198  
Dh 3749 tatcatcatgtatagttctgagtggaagaagtggtgcaagactatctgagaagaattccacttg 3808  
QY 4199 aaccttaatgtctctcttcttgcaatgttgatccaactctggataaacaacataatttttaa 4258  
Dh 3809 aaccttaatgtctctcttcttgcaatgttgatccaactctggataaacaacataatttttaa 3868

QY 4259 cctccaaatatttctgtgatatacatccgtttgcccacaaactatratatgacttgcagcac 4318  
Dh 3869 cctccaaatatttctgtgatatacatccgtttgcccacaaactatratatgacttgcagcac 3928  
QY 4319 tcttcgcatgtgagttgagtggtctggtatataagtttgacagatccatctgggaatgga 4378  
Dh 3929 tcttcgcatgtgagttgagtggtctggtatataagtttgacagatccatctgggaatgga 3988  
QY 4379 gagtaagcaatatacagatgacagatbaactgtcttcaacttcaacttcaactatgtttgc 4438  
Dh 3989 gagtaagcaatatacagatgacagatbaactgtcttcaacttcaacttcaactatgtttgc 4048  
QY 4439 cacctgtctctccttcaaaagctgacatcttcaacttcaaaaggagagatgtgtgagacc 4498  
Dh 4049 cacctgtctctccttcaaaagctgacatcttcaacttcaaaaggagagatgtgtgagacc 4108  
QY 4499 tcaagtgataataatcaaaaggtgtgctgcaagtggaacttccagaagaacaatgaaatgcaac 4558  
Dh 4109 tcaagtgataataatcaaaaggtgtgctgcaagtggaacttccagaagaacaatgaaatgcaac 4168  
QY 4559 aggaagtaactactccaaggtgtaaaatctctgtcttaccagcaatgatatgaaagattcct 4618  
Dh 4169 aggaagtaactactccaaggtgtaaaatctctgtcttaccagcaatgatatgaaagattcct 4228  
QY 4619 catctccagcgtaagaatggtgcaatcagtgagactctcttcttcaagatggcaagaatgaa 4678  
Dh 4229 catctccagcgtaagaatggtgcaatcagtgagactctcttcttcaagatggcaagaatgaa 4288  
QY 4679 gtttttccaggaataatcaagctccttcaacacctgtgtgtaactctcttcaagccacaggt 4738  
Dh 4289 gtttttccaggaataatcaagctccttcaacacctgtgtgtaactctcttcaagccacaggt 4348  
QY 4739 actgaactgtacacttcaagatcaacccccagagttggtgacacagatctgacctgaagat 4798  
Dh 4349 actgaactgtacacttcaagatcaacccccagagttggtgacacagatctgacctgaagat 4408  
QY 4799 ggaagttctgtggtcgagagcgacaggaactctactgagaggtgagcaatc 4848  
Dh 4409 ggaagttctgtggtcgagagcgacaggaactctactgagaggtgagcaatc 4458

RESULT 10  
AAA49232  
ID AAA49232 standard; DNA: 12445 BP.  
XX  
AC AAA49232;  
XX  
DT 26-SEP-2000 (first entry)  
XX  
DE Vector HSReneo for transforming endothelial cells.  
XX  
XX Hemostatic; antianemic; antidiabetic; nootropic; neuroprotective; vector;  
KW osteopathic; antislackling; immunostimulant; gene therapy; collagen;  
KW endothelial cell; peripheral blood; buffy coat cell; VEGF; PCR primer;  
KW vascular endothelial growth factor; bovine brain extract; haemophilla;  
KW Factor VIII; human; transgene; adenosine deaminase deficiency; ss;  
KW sickle cell anaemia; thalassemia; diabetes; alpha-antitrypsin deficiency;  
KW Alzheimer's disease; brain disease; heart disease; immune system defect;  
KW bone fracture; osteoporosis.  
OS Homo sapiens.  
OS Synthetic.  
XX  
XX WO200032750-A1.  
XX  
XX 08-JUN-2000.  
XX  
XX PD  
XX PF 24-NOV-1999; 99MO-US28033.  
XX  
XX PR 24-NOV-1998; 98US-0109687.  
XX  
XX PA (MIND ) UNIV MINNESOTA.



| Query Match           | 54.4%;       | Score 4319;  | DB 21;     | Length 12445;      |
|-----------------------|--------------|--|------------|--------------------|
| Best Local Similarity | 98.9%;       | Pred. No. 0;   |            |                    |
| Matches 4374;         | Conservative | 0;   | Mismatches | 5; Indels 42; Gaps |
| 0Y                    | 415          | ccaccatgcaaatagagctctccaccctgcttcttcttgcttgcttttgagatctgctta   | 474        |                    |
| Db                    | 763          | ccaccatgcaaatagagctctccaccctgcttcttcttgcttgcttttgagatctgctta   | 822        |                    |
| 0Y                    | 475          | gtgcaccgcaagaatactactccttggtgagctggaactgtcactgtgagctatgtcaaatg | 534        |                    |
| Db                    | 823          | gtgcaccgcaagaatactactccttggtgagctggaactgtcactgtgagctatgtcaaatg | 882        |                    |
| 0Y                    | 535          | atctcgtgtagctgctgtggaagcaaatcttctctctagatgccaataatctttccat     | 594        |                    |
| Db                    | 883          | atctcgtgtagctgctgtggaagcaaatcttctctctagatgccaataatctttccat     | 942        |                    |
| 0Y                    | 595          | tcaaacccctcagtcgtgtaaaaaaacctctgtttgaatatcaaggttcaacctttta     | 654        |                    |
| Db                    | 943          | tcaaacccctcagtcgtgtaaaaaaacctctgtttgaatatcaaggttcaacctttta     | 1002       |                    |
| 0Y                    | 655          | acatcgctaaagcaagccacacctgtgattggttctgtaagctctacacaccagctgag    | 714        |                    |
| Db                    | 1003         | acatcgctaaagcaagccacacctgtgattggttctgtaagctctacacaccagctgag    | 1062       |                    |
| 0Y                    | 715          | tttatgatacagtggtcatctaacactaaagaacatggtcttccactctgtcagtctcatg  | 774        |                    |
| Db                    | 1063         | tttatgatacagtggtcatctaacactaaagaacatggtcttccactctgtcagtctcatg  | 1122       |                    |
| 0Y                    | 775          | cgtgtgtgtatccctactggaagctcttcgaggaagctgaatatgattgttcagcaagtc   | 834        |                    |
| Db                    | 1123         | cgtgtgtgtatccctactggaagctcttcgaggaagctgaatatgattgttcagcaagtc   | 1182       |                    |
| 0Y                    | 835          | aaagggagaaagaatgataagctcttccctgtgtggaagcacaataatgtctggaag      | 894        |                    |

|    |      |   |      |
|----|------|---|------|
| Db | 1183 | aaagggaagaaagatgataaagtctccctgvtggaagcaataatgvtcgcag              | 124  |
| Qy | 895  | tcttgaagagaatgvtccaatggcctctgaaccaatgvtccttaacctatactt            | 954  |
| Db | 1243 | tcttgaagagaatgvtccaatggcctctgaaccaatgvtccttaacctatactt            | 1302 |
| Qy | 955  | ctcatgtgacctgtttaaagacttgatctcagccctatggagccctacagatga            | 1014 |
| Db | 1303 | ctcatgtgacctgtttaaagacttgatctcagccctatggagccctacagatga            | 1362 |
| Qy | 1015 | gagaaaggagctctgcaccaagaaaagacaagaccttgcaaaaattataactcttg          | 1074 |
| Db | 1363 | gagaaaggagctctgcaccaagaaaagcacagaccttgcaaaaattataactcttg          | 1422 |
| Qy | 1075 | ctgtatttgatgaaagggaagaatttggaactcagaacaagaactccttgatgcagata       | 1133 |
| Db | 1423 | ctgtatttgatgaaagggaagaatttggaactcagaacaagaactccttgatgcagata       | 1482 |
| Qy | 1135 | ggagatgcgactctgcctcgcgcgcgtcaaaatgcaacagatgaatggttatgaaca         | 1194 |
| Db | 1483 | ggagatgcgactctgcctcgcgcgcgtcaaaatgcaacagatgaatggttatgaaca         | 1544 |
| Qy | 1195 | ggctctgcgcaagtcgtatggaatgcacagaaatcgaatctatggaatggaatgga          | 1254 |
| Db | 1543 | ggctctgcgcaagtcgtatggaatgcacagaaatcgaatctatggaatggaatgga          | 1602 |
| Qy | 1255 | tgggacacacccctggaaagtgcaccaatactcctcgaaaggtcaacacattctctggaga     | 1314 |
| Db | 1603 | tgggacacacccctcggaagtgcaccaatactcctcgaaaggtcaacacattctctggaga     | 1662 |
| Qy | 1315 | accatcgcagagcgvtccttggaaatctcgcgaataactcttcctactgtcaaacctc        | 1374 |
| Db | 1663 | accatcgcagagcgvtccttggaaatctcgcgaataactcttcctactgtcaaacctc        | 1722 |
| Qy | 1375 | tgaatgaccttggaacgtttctactgttttggcaactcttcgcccaacagtgtgca          | 1434 |
| Db | 1723 | tgaatgaccttggaacgtttctactgttttggcaactcttcgcccaacagtgtgca          | 1782 |
| Qy | 1435 | tggaagccttatgtccaagttagacacgctgtccagaagaaacccaactacgaatgaanaata   | 1494 |
| Db | 1783 | tggaagccttatgtccaagttagacacgctgtccagaagaaacccaactacgaatgaanaata   | 1842 |
| Qy | 1495 | atgaagaagcgggaagacatgaatgatacttactgattctgaatagatgvtgvtcgaagt      | 1555 |
| Db | 1843 | atgaagaagcgggaagacatgaatgatacttactgattctgaatagatgvtgvtcgaagt      | 1902 |
| Qy | 1555 | ttgaatgaagaaactctctctcccttataccaatctcgctcagttgtccaagaagcatcccta   | 1614 |
| Db | 1903 | ttgaatgaagaaactctctctcccttataccaatctcgctcagttgtccaagaagcatcccta   | 1962 |
| Qy | 1615 | aaacttgggtacatatacttgctgctgtaagaaggaagacttgggaactatgctcccttaagtc  | 1674 |
| Db | 1963 | aaacttgggtacatatacttgctgctgtaagaaggaagacttgggaactatgctcccttaagtc  | 2022 |
| Qy | 1675 | tcgcccccgatgaagaaagtataaagaatcaatatttggacaatggccctcagcgagatg      | 1734 |
| Db | 2023 | tcgcccccgatgaagaaagtataaagaatcaatatttggacaatggccctcagcgagatg      | 2082 |
| Qy | 1735 | gtagaagaatatacaaaaagctccgaatttatggcatatacagaatgaanaaccttaagactctg | 1794 |
| Db | 2083 | gtagaagaatatacaaaaagctccgaatttatggcatatacagaatgaanaaccttaagactctg | 2144 |
| Qy | 1795 | aagcatctcagaatcgaaatcgaatctctgggaaccttacttatatgvggaagtltggaaga    | 1855 |
| Db | 2143 | aagcatctcagaatcgaaatcgaatctctgggaaccttacttatatgvggaagtltggaaga    | 2202 |
| Qy | 1855 | caacttgaattatatttaagaatacaagcaagcagacataataacatctacccctcagga      | 1914 |
| Db | 2203 | caacttgaattatatttaagaatacaagcaagcagacataataacatctacccctcagga      | 2262 |
| Qy | 1915 | tcaactgaatgcgtcctcttgatccaaggagatatacaaaaagvtgtaaaaacttgaagg      | 1974 |
| Db | 2263 | tcaactgaatgcgtcctcttgatccaaggagatatacaaaaagvtgtaaaaacttgaagg      | 2322 |



QY 1975 atttccaatctgcgaaggagaaatattcaaatataatgagacgtgacgtgagaagt 2034  
|||||  
Db 2323 atttccaatctgcgcgaaggagaaatattcaaatataatgagacgtgacgtgagaagt 2382  
|||||  
QY 2035 ggccaataaatacgaatccctgcgtgcgtgcacgcgtattacttagtttcgttaaatg 2094  
|||||  
Db 2383 ggccaataaatacgaatccctgcgtgcgtgcacgcgtattacttagtttcgttaaatg 2442  
|||||  
QY 2095 agagagatctgcttcgagactcaattggcctctccctcattctgtctacaaaagactgtag 2134  
|||||  
Db 2443 agagagatctgcttcgagactcaattggcctctccctcattctgtctacaaaagactgtag 2502  
|||||  
QY 2155 atcaaaaggagaaacagaataatgtaagacaagagaagtgtcatcctgttctctgatttg 2214  
|||||  
Db 2503 atcaaaaggagaaacagaataatgtaagacaagagaagtgtcatcctgttctctgatttg 2562  
|||||  
QY 2215 atgagaacgcgaagctgtgtaccctcacagagaaatatacaagcttctccccaatccagctg 2274  
|||||  
Db 2563 atgagaacgcgaagctgtgtaccctcacagagaaatatacaagcttctccccaatccagctg 2622  
|||||  
QY 2275 gagtgcagcttgaggaatcccaagtcccaagctccacaataatgcacagcatcaatgct 2334  
|||||  
Db 2623 gagtgcagcttgaggaatcccaagtcccaagctccacaataatgcacagcatcaatgct 2682  
|||||  
QY 2335 atgttttgatagtttgcaagttgcaagttgtttgcatgagtgagcatggtacattc 2394  
|||||  
Db 2683 atgttttgatagtttgcaagttgcaagttgtttgcatgagtgagcatggtacattc 2742  
|||||  
QY 2395 taagcatctggaacacagactgactcctctctctctctctctctctctctctctctctcaac 2454  
|||||  
Db 2743 taagcatctggaacacagactgactcctctctctctctctctctctctctctctctctcaac 2802  
|||||  
QY 2455 acaaaatgtgtatagaagacacactcaaccctattccattctcagaggaacgtctctca 2514  
|||||  
Db 2803 acaaaatgtgtatagaagacacactcaaccctattccattctcagaggaacgtctctca 2862  
|||||  
QY 2515 tgtcatgtaaaacccaggtctatggtctctgggtgcgaacactccagaacttcggaac 2574  
|||||  
Db 2863 tgtcatgtaaaacccaggtctatggtctctgggtgcgaacactccagaacttcggaac 2922  
|||||  
QY 2575 gaggcatagcgcgccttaactgaagttctctagttgtgaacaagaacactgtgtattacg 2634  
|||||  
Db 2923 gaggcatagcgcgccttaactgaagttctctagttgtgaacaagaacactgtgtattacg 2982  
|||||  
QY 2635 aggcagttatagaagatatattcagcatactgtcgtgataaacaatgtgccattgaacaa 2694  
|||||  
Db 2983 aggcagttatagaagatatattcagcatactgtcgtgataaacaatgtgccattgaacaa 3042  
|||||  
QY 2695 gaagcttctccagaattcaagacacccctagcactagcaaaagcaatttaatgcaccc 2754  
|||||  
Db 3043 gaagcttctccagaattcaagacacccctagcactagcaaaagcaatttaatgcaccc 3060  
|||||  
QY 2755 caccagttctgaaacgcgcatacgcggaataaactcgtactactctcagtcagatacag 2814  
|||||  
Db 3061 caccagttctgaaacgcgcatacgcggaataaactcgtactactctcagtcagatacag 3120  
|||||  
QY 2815 aggaatattgataatgataccatattcagttgaaatgaaagaagaattttgaattt 2874  
|||||  
Db 3121 aggaatattgataatgataccatattcagttgaaatgaaagaagaattttgaattt 3180  
|||||  
QY 2875 atgataaggaatgaataatcagaagcccgcaagcttcaaaaagaacgcgacacatttta 2934  
|||||  
Db 3181 atgataaggaatgaataatcagaagcccgcaagcttcaaaaagaacgcgacacatttta 3240  
|||||  
QY 2935 ttgctgcagtgagagagcctcgtgattatagatgagtagctccccaatgttctaagaa 2994  
|||||  
Db 3241 ttgctgcagtgagagagcctcgtgattatagatgagtagctccccaatgttctaagaa 3300  
|||||  
QY 2995 acaaggtctcagagtgagcagtgctccctcagttcaagaagaagtgttttccagaatttactg 3054  
|||||  
Db 3301 acaaggtctcagagtgagcagtgctccctcagttcaagaagaagtgttttccagaatttactg 3360  
|||||

QY 3055 atgctccttactcagcccttataccgtgagagaactaaatgaacatttggactcctg 3114  
|||||  
Db 3361 atgctccttactcagcccttataccgtgagagaactaaatgaacatttggactcctg 3420  
|||||  
QY 3115 ggccataataagacagaagtctgaagataatcatgtgtaacttccagaatacaggct 3174  
|||||  
Db 3421 ggccataataagacagaagtctgaagataatcatgtgtaacttccagaatacaggct 3480  
|||||  
QY 3175 ctgctccctattct 3234  
|||||  
Db 3481 ctgctccctattct 3500  
|||||  
QY 3235 cagaacctagaanaaacttctcaagcctaaatgaacccaacttacttctggaagtgc 3294  
|||||  
Db 3541 cagaacctagaanaaacttctcaagcctaaatgaacccaacttacttctggaagtgc 3600  
|||||  
QY 3295 aacatcatgccaacccaactaaagatgagtttgactgcgaagcctggtctattctctg 3354  
|||||  
Db 3601 aacatcatgccaacccaactaaagatgagtttgactgcgaagcctggtctattctctg 3660  
|||||  
QY 3355 atgttgacctggaanaagatgtgcactcagcctgatttggaacccctctgtctgcaca 3414  
|||||  
Db 3661 atgttgacctggaanaagatgtgcactcagcctgatttggaacccctctgtctgcaca 3720  
|||||  
QY 3415 ctaacacactgaacccctgctcatgtgagacagtgacagtaacgaatttgcctgttt 3474  
|||||  
Db 3721 ctaacacactgaacccctgctcatgtgagacagtgacagtaacgaatttgcctgttt 3780  
|||||  
QY 3475 tcaacatcttgatgagagccaagaagctgtgtacttcaactgaaataatggaagaactgca 3534  
|||||  
Db 3781 tcaacatcttgatgagagccaagaagctgtgtacttcaactgaaataatggaagaactgca 3840  
|||||  
QY 3535 gggtcccttgataatatacagtgagaatgccacttttaagaagaattatgtcttcaatg 3594  
|||||  
Db 3841 gggtcccttgataatatacagtgagaatgccacttttaagaagaattatgtcttcaatg 3900  
|||||  
QY 3595 caatcaatggtctataataatgatacaactcctgtctgataatgtgctcagagatacaaga 3654  
|||||  
Db 3901 caatcaatggtctataataatgatacaactcctgtctgataatgtgctcagagatacaaga 3960  
|||||  
QY 3655 ttgcatgtgatactgcctcagcatggtgcagcaatgaanaatccattctatctcaattcagtg 3714  
|||||  
Db 3961 ttgcatgtgatactgcctcagcatggtgcagcaatgaanaatccattctatctcaattcagtg 4020  
|||||  
QY 3715 gacatgtgtactcgttgcgaanaaagaagaggttaaatgtgcctgtacatctctatc 3774  
|||||  
Db 4021 gacatgtgtactcgttgcgaanaaagaagaggttaaatgtgcctgtacatctctatc 4080  
|||||  
QY 3775 cagggtttttgagacagtggaatgtttacatccaagaactggaatttggcgggtggaat 3834  
|||||  
Db 4081 cagggtttttgagacagtggaatgtttacatccaagaactggaatttggcgggtggaat 4140  
|||||  
QY 3835 gccattatggcagcatctacatgctcgtgagatgagacactttctcgtgtacagcaata 3894  
|||||  
Db 4141 gccattatggcagcatctacatgctcgtgagatgagacactttctcgtgtacagcaata 4200  
|||||  
QY 3895 agtctcagactccctcgtggaatggtctctgcgacacattagaagattttcagattacgctt 3954  
|||||  
Db 4201 agtctcagactccctcgtggaatggtctctgcgacacattagaagattttcagattacgctt 4260  
|||||  
QY 3955 caggaacatatgagacagtggtgcccccaagctgcgcagactcatattccgcgataca 4014  
|||||  
Db 4261 caggaacatatgagacagtggtgcccccaagctgcgcagactcatattccgcgataca 4320  
|||||  
QY 4015 atgcctggaagcaaccaagagcccttctctgagatcaaggtgagatctgttgccaccaatga 4074  
|||||  
Db 4321 atgcctggaagcaaccaagagcccttctctgagatcaaggtgagatctgttgccaccaatga 4380  
|||||  
QY 4075 ttattcaggaatcaagaacccaggtgtcccggtgagaagttctcagagctctacatctctc 4134  
|||||  
Db 4381 ttattcaggaatcaagaacccaggtgtcccggtgagaagttctcagagctctacatctctc 4440  
|||||  
QY 4135 agttatcatcatgttatagttctgatgtggaagaagtgtgcagacttatcgaagaattcca 4194  
|||||

|           |   |  |      |
|-----------|---|--|------|
| Db        | 4441  | agttattacacagtatagctctgagtggaagaagtggcagacttacgcgaataacc       | 4500 |
| Qy        | 4195  | ctggaaccttaagtgctctctcttgccaatgtggatcatctgggataaacaacata       | 4254 |
| Db        | 4501  | ctggaaccttaagtgctctctcttgccaatgtggatcatctgggataaacaacata       | 4560 |
| Qy        | 4255  | ttaacctccaattatctgctcgatactccgtttgcacccaactcatatagcttcga       | 4314 |
| Db        | 4561  | ttaacctccaattatctgctcgatactccgtttgcacccaactcatatagcttcga       | 4620 |
| Qy        | 4315  | gcactctcgacgagtgatgagcgctgatataaatgttgacagatgcatctggaa         | 4374 |
| Db        | 4621  | gcactctcgacgagtgatgagcgctgatataaatgttgacagatgcatctggaa         | 4680 |
| Qy        | 4375  | tggagagtaaaagcaatacagatgacagatctactgtcatccctacttaaccaat        | 4434 |
| Db        | 4681  | tggagagtaaaagcaatacagatgacagatctactgtcatccctacttaaccaat        | 4740 |
| Qy        | 4435  | tgcacacgtgctctccttcaaaaagctcgacttaacctccaaggagagatgacctga      | 4494 |
| Db        | 4741  | tgcacacgtgctctccttcaaaaagctcgacttaacctccaaggagagatgacctga      | 4800 |
| Qy        | 4495  | gacctagatgataatacacaagaatggtctgcagaatgcatccacagaagaacatga      | 4554 |
| Db        | 4801  | gacctagatgataatacacaagaatggtctgcagaatgcatccacagaagaacatga      | 4860 |
| Qy        | 4555  | tccaaagagtaactactccaaggagtaaatctctgcttaccagatgatgtgaagagt      | 4614 |
| Db        | 4861  | tccaaagagtaactactccaaggagtaaatctctgcttaccagatgatgtgaagagt      | 4920 |
| Qy        | 4615  | tccctactccaagcagtcacaagatggccatagtgagactctcttctccaagaatgcaag   | 4674 |
| Db        | 4921  | tccctactccaagcagtcacaagatggccatagtgagactctcttctccaagaatgcaag   | 4980 |
| Qy        | 4675  | taaaggttttcagaaggaaatccaagaactcctctcacacctgtgtggaactcttagaccac | 4734 |
| Db        | 4981  | taaaggttttcagaaggaaatccaagaactcctctcacacctgtgtggaactcttagaccac | 5040 |
| Qy        | 4735  | cgtactgactgcctacactcgatctacccccagatgtggatgcacagatgcctga        | 4794 |
| Db        | 5041  | cgtactgactgcctacactcgatctacccccagatgtggatgcacagatgcctga        | 5100 |
| Qy        | 4795  | gattggaagttctgggctgcgagacagagaaactctactga 4835                 |      |
| Db        | 5101  | gattggaagttctgggctgcgagacagagaaactctactga 5141                 |      |
| RESULT 11 |   |  |      |
| AAN90654  |   |  |      |
| ID        | AAN90654  | standard; DNA; 4275 BP.  |      |
| AC        | AAN90654;   |  |      |
| XX        |   |  |      |
| DT        | 26-JUN-1990   | (first entry)  |      |
| XX        |   |  |      |
| DE        | DNA encoding 740 Arg-1649 Glu human Factor VIII.C.              |  |      |
| XX        |   |  |      |
| KW        | Human Factor VIII.C; Ad.RE.neo; 740 Arg-1649 Glu Factor VIII.C; |  |      |
| KW        | haemophilia A.  |  |      |
| XX        |   |  |      |
| OS        | Homo sapiens.   |  |      |
| XX        |   |  |      |
| FH        | Key   | Location/Qualifiers  |      |
| FT        | CDS   | 1..4275  |      |
| FT        |   | /tag= a  |      |
| XX        |   |  |      |
| PN        | EP306968-A.   |  |      |
| XX        |   |  |      |
| PD        | 15-MAR-1989.  |  |      |
| XX        |   |  |      |
| PF        | 09-SEP-1988;  | 88EP-0114769.  |      |
| XX        |   |  |      |

PR 08-APR-1988; 88JP-0085454.  
XX  
XX (KAGA ) CHEMO-SERO-THERAP.  
PA (TEIJ ) TEIJIN LTD.  
XX  
PI Sugiyama T, Masuda K, Tajima Y, Yonemura H;  
XX  
XX WPI; 1989-078467/11.  
DR P-PSDB; AAP91165.  
XX  
PT Prodn. of recombinant human Factor-VIII-C -  
PT using animal cells transformed with a vector contg. the gene for  
PT Factor VIII:C and a promoter  
XX  
XX  
PS Fig 1(1) - 1(13); ; 32pp; English.  
XX  
XX When translated, Arg-740 of the carboxyl terminus of the H chain is  
CC directly bonded by a peptide bond to Glu-1649 of the amino terminus of  
CC L chain. It is used to transform animal cells so that they produce  
CC human Factor VIII:C. A pref. expression vector is plasmid Ad. RE. neo.  
CC The expression vector has at least one promoter upstream of AAN90654.  
CC The transformants can constantly and continuously produce human Factor  
CC VIII:C in high yield on a commercial scale. The human Factor VIII:C so  
CC produced is considered to corresp. to the smallest species of active and  
CC intact Factor VIII:C molecules in the human blood plasma. It is useful  
CC for treating haemophilia A patients.  
XX  
XX Sequence 4275 BP; 1245 A; 941 C; 945 G; 1144 T; 0 other;  
SQ

|                            |        |  |        |                 |
|----------------------------|--------|--|--------|-----------------|
| Query Match                | 52.5%  | Score 4173;  | DB 10; | Length 4275;    |
| Best Local Similarity      | 96.0%; | Pred. No. 0;   |        |                 |
| Matches 4270; Conservative | 0;     | Mismatches   | 5;     | Indels 84; Gaps |
| OY                         | 477    | gccaccagaatactactctggtgcagctggaactgcatatgagactatgcacaagatg       | 536    |                 |
|                            |        |  |        |                 |
| Db                         | 1      | gccaccagaatactactactctggtgcagctggaactgcatatgagactatgcacaagatg    | 60     |                 |
| OY                         | 537    | ctcggctgagctgcctctgtagacgcaagattctccctcagtagtgcacaaatctttccatc   | 596    |                 |
|                            |        |  |        |                 |
| Db                         | 61     | ctcggctgagctgcctctgtagacgcaagattctccctcagtagtgcacaaatctttccatc   | 120    |                 |
| OY                         | 597    | aacacctgaatcgcgtgtacaaaaaagactctgttttagaatccaaggttcaaccttttcaac  | 656    |                 |
|                            |        |  |        |                 |
| Db                         | 121    | aacacctgaatcgcgtgtacaaaaaagactctgttttagaatccaaggttcaaccttttcaac  | 180    |                 |
| OY                         | 657    | atcgctaaagccaagccacccctgtagctggtctctgtagtctccacacccaagctgaagt    | 716    |                 |
| Db                         | 161    | atcgctaaagccaagccacccctgtagtctggtctctgtagtctccacacccaagctgaagt   | 240    |                 |
| OY                         | 717    | tatgatacagctggttcattacacactaaagaacatgagctcccatcctcgtcaatctcatgct | 776    |                 |
|                            |        |  |        |                 |
| Db                         | 241    | tatgatacagctggttcattacacactaaagaacatgagctcccatcctcgtcaatctcatgct | 300    |                 |
| OY                         | 777    | gttggctgatactctactgtgaaagctcttcgaaggagctgaataatgtagtgcagaacagca  | 836    |                 |
| Db                         | 301    | gttggctgatactctactgtgaaagctcttcgaaggagctgaataatgtagtgcagaacagca  | 360    |                 |
| OY                         | 837    | aggaggaagaagaatgatataaagctctccctctgtggaagccataaataatgtctggcagtc  | 896    |                 |
|                            |        |  |        |                 |
| Db                         | 361    | aggaggaagaagaatgatataaagctctccctctgtggaagccataaataatgtctggcagtc  | 420    |                 |
| OY                         | 897    | ctgaaagagaatggtccaaatggtcctctgaccccaactgtgccttacctatcatattctc    | 956    |                 |
|                            |        |  |        |                 |
| Db                         | 421    | ctgaaagagaatggtccaaatggtcctctgaccccaactgtgccttacctatcatattctc    | 480    |                 |
| OY                         | 957    | catgtggagcctgggtaaaagacttaattcaagagctcaattgagccctacagtatgaga     | 1016   |                 |
|                            |        |  |        |                 |
| Db                         | 481    | catgtggagcctgggtaaaagacttaattcaagagctcaattgagccctacagtatgaga     | 540    |                 |
| OY                         | 1017   | gaaggagatctctgccaagaaagacacagaacctgtgacaaattatatactatttgcct      | 1076   |                 |
|                            |        |  |        |                 |
| Db                         | 541    | gaaggagatctctgccaagaaagacacagaacctgtgacaaattatatactatttgcct      | 600    |                 |

QY 1077 gtatttgatgaagggaagaagtgtgacatcagaaaacaaagaactccttgatcagatagg 1136  
|||||  
Db 601 gtaatttgatgaagggaagaagtttggcaaccagaaagaactcccttgatgacagatagg 660  
QY 1137 gatgctgcatcgtcgcggccttggtcctaaatgacacagccaatggtgtatgtaaacagg 1196  
|||||  
Db 661 gatgctgcatcgtcgcggccttggtcctaaatgacacagccaatggtgtatgtaaacagg 720  
QY 1197 tctctgcgaagttcgtatgtgtatgacgaagaatcagttctatggtcatgtgtatggaaatg 1256  
|||||  
Db 721 tctctaccaagttcgtatgtgtatgacgaagaatcagttctatggtcatgtgtatggaaatg 780  
QY 1257 ggaacacactcctggaagtgcaatcaatatctctgaaagtgacacatcttctgttgaagac 1316  
|||||  
Db 781 ggaacacactcctggaagtgcaatcaatatctctgaaagtgacacatcttctgttgaagac 840  
QY 1317 catgcgcagagcgtccttggaatatctgcgcaataacttctctactgtctcaaacctcttg 1376  
|||||  
Db 841 catgcgcagagcgtccttggaatatctgcgcaataacttctctactgtctcaaacctcttg 900  
QY 1377 atggaacttggaacagtttctactgttttgtatatactctctcccaacatgtagcatg 1436  
|||||  
Db 901 atggaacttggaacagtttctactgttttgtatatactctctcccaacatgtagcatg 960  
QY 1437 gaagcttatgtccaagtagagcagctgtccagagaagaccccaactagcaatgaaataat 1496  
|||||  
Db 961 gaagcttatgtccaagtagagcagctgtccagagaagaccccaactagcaatgaaataat 1020  
QY 1497 gaaagaacggaagaactatgatagcatcttactatctcgaatgtagtgcagttt 1556  
|||||  
Db 1021 gaagaacggaagaactatgatagcatcttactatctcgaatgtagtgcagttt 1080  
QY 1557 gatgtatgaacaactcctctctctatccaatctcgctcaagtgtgcgaagaagcatccta 1616  
|||||  
Db 1081 gatgatgaacaactcctctctctatccaatctcgctcaagtgtgcgaagaagcatccta 1140  
QY 1617 actgtggttacatctatctgctgcgtgaagaagtagagcttggtcctcttgctctc 1676  
|||||  
Db 1141 actgtggttacatctatctgctgcgtgaagaagtagagcttggtcctcttgctctc 1200  
QY 1677 gcccccagatgacagaagttataaagaatcaatatcttgacaatggtccctcagcggatggt 1736  
|||||  
Db 1201 gcccccagatgacagaagttataaagaatcaatatcttgacaatggtccctcagcggatggt 1260  
QY 1737 aggaagaatcaaaaaagtcggatcttatgtgcataacagatgaaaccttgaacctcgtaa 1796  
|||||  
Db 1261 aggaagaatcaaaaaagtcggatcttatgtgcataacagatgaaaccttgaacctcgtaa 1320  
QY 1797 gctattcaagcatgacaggaatcttggaaccttactttagtgggaagttggagacaca 1856  
|||||  
Db 1321 gctattcaagcatgacaggaatcttggaaccttactttagtgggaagttggagacaca 1380  
QY 1857 ctgttgatattatattaaagaatcaagaacagacacataaactaccctcaagcatc 1916  
|||||  
Db 1381 ctgttgatattatattaaagaatcaagaacagacacataaactaccctcaagcatc 1440  
QY 1917 actgagtctcgctccttctgtatccaagaagattaccaaaagtgtaaaacatttgaagat 1976  
|||||  
Db 1441 actgagtctcgctccttctgtatccaagaagattaccaaaagtgtaaaacatttgaagat 1500  
QY 1977 ttccaattctgcgcagaggaataatccaataatgaagtagagctgtgaagaatggg 2036  
|||||  
Db 1501 ttccaattctgcgcagaggaataatccaataatgaagtagagctgtgaagaatggg 1560  
QY 2037 ccaactaatcagatccctcggtgcctgacccgctattacttagtttcgttaatatggag 2096  
1561 ccaactaatcagatccctcggtgcctgacccgctattacttagtttcgttaatatggag 1620  
QY 2097 agagaatcagttcaggaactaatgtgcccctcctcctcatctgctacaagaatcgttagat 2156  
|||||  
Db 1621 agagaatcagttcaggaactaatgtgcccctcctcctcatctgctacaagaatcgttagat 1680

QY 2157 caaagaggaaccagataatgtccagacaaggaagatgtcatcctgtttctgtatttgat 2216  
|||||  
Db 1681 caaagaggaaccagataatgtccagacaaggaagatgtcatcctgtttctgtatttgat 1740  
QY 2217 gagaacccaagctgtgtactccagagaataataacagcttcttcccaatccagctgtga 2276  
1741 gagaacccaagctgtgtactccagagaataataacagcttcttcccaatccagctgtga 1800  
QY 2277 gtgcagcttgagatccagagttccaaagcctccacaatcatgtacagacataatggtcat 2336  
1801 gtgcagcttgagatccagagttccaaagcctccacaatcatgtacagacataatggtcat 1860  
QY 2337 gttcttgaatgttcgaatttcagttgtgttgcataaggtgggacatactgttacaacta 2396  
1861 gttcttgaatgttcgaatttcagttgtgttgcataaggtgggacatactgttacaacta 1920  
QY 2397 agcatgtgagcacagactgtactcctctctgtctctctctgtgatatacccttcaaacac 2456  
1921 agcatgtgagcacagactgtactcctctctgtctctctctgtgatatacccttcaaacac 1980  
QY 2457 aaaaatgtctatgaaagaacacactaacctatcccatctcaagagaacagctctcatg 2516  
1981 aaaaatgtctatgaaagaacacactaacctatcccatctcaagagaacagctctcatg 2040  
QY 2517 tcgattggaaaacccaaggtctatgtgattctgggtgcgcaacactagactttcggaaaca 2576  
2041 tcgattggaaaacccaaggtctatgtgattctgggtgcgcaacactagactttcggaaaca 2100  
QY 2577 ggcattgaacgccttacttgaagtttcttagttgtgtacaaagaacactgtgtattatcagag 2636  
2101 ggcattgaacgccttacttgaagtttcttagttgtgtacaaagaacactgtgtattatcagag 2160  
QY 2637 gacagttatgaaatatttcaagcatctgctgtggttaaaaaaatgcatatgaaacaa 2696  
2161 gacagttatgaaatatttcaagcatctgctgtggttaaaaaaatgcatatgaaacaa 2220  
QY 2697 agcttctccagaattcaagcacccctagcactagcgaaagaacatattatgcaaccca 2756  
2221 ----- 2220  
QY 2757 ccaagtttgaaaacccaacaaaggaataactcgttactcttcaagtcagatcaagaag 2816  
2221 ----- gaaataactcgttactcttcaagtcagatcaagaag 2256  
QY 2817 gaaattgactatgatatcaacatatccgttgaatgaagaaggaagatttgaacttat 2876  
2257 gaaattgactatgatatcaacatatccgttgaatgaagaaggaagatttgaacttat 2316  
QY 2877 gatgagatgaaatacagagcccccagcttccaagaagaacacagacatatattat 2936  
2317 gatgagatgaaatacagagcccccagcttccaagaagaacacagacatatattat 2376  
QY 2937 gctgcagtgagaggtctcgtggaattatggaatgagtagctcccaatgttctaagaac 2996  
2377 gctgcagtgagaggtctcgtggaattatggaatgagtagctcccaatgttctaagaac 2436  
QY 2997 agggctcaagtggtgcagttccctcagttccaagaagttgtttcccaagatttactgat 3056  
2437 agggctcaagtggtgcagttccctcagttccaagaagttgtttcccaagatttactgat 2486  
QY 3057 ggtcccttactcagcccttataccgttggagaagaataatgaaacttgggactcctgggg 3116  
2497 ggtcccttactcagcccttataccgttggagaagaataatgaaacttgggactcctgggg 2556  
QY 3117 ccaataataagagcagaagttgaaagaataataatcattgtaaatctcagaatccagcctct 3176  
2557 ccaataataagagcagaagttgaaagaataataatcattgtaaatctcagaatccagcctct 2616  
QY 3177 cgtccctattctcctattctcagccttatcttcatatgaggaagatcagaagcagaagca 3236  
2617 cgtccctattctcctattctcagccttatcttcatatgaggaagatcagaagcagaagca 2676  
QY 3237 gaacctagaaaaaacttctgtaagcctaattgaaaccaaacttacttcttggaagttgaa 3296

Db 2677 gaacctgaaaaaactctgtcaagcctaaatgaaaccaaactactcttctggaatgcaaa 2736  
Qy 3297 catcatatgagccaccactaaagaatgattctgacgcaaaagccttggtctattctctgat 3356  
Db 2737 catcatatgagccaccactaaagaatgattctgacgcaaaagccttggtctattctctgat 2796  
Qy 3357 gttagacctggaanaaagatgtgcaactcaaggcctgatatggaacccctctgtctgcacact 3416  
Db 2797 gttagacctggaanaaagatgtgcaactcaaggcctgatatggaacccctctgtctgcacact 2856  
Qy 3417 aacacactgaaacctgtctcattgaggagacaagtgtacatcagagaattgtctgtcttc 3476  
Db 2857 aacacactgaaacctgtctcattgaggagacaagtgtacatcagagaattgtctgtcttc 2916  
Qy 3477 accactcttgtaggaacaaagaagcgtgactcctcactgtaaaatataatgaaagaacgtcag 3536  
Db 2917 accactcttgtaggaacaaagaagcgtgactcctcactgtaaaatataatgaaagaacgtcag 2976  
Qy 3537 gctccctgcaaatatccagatggaagaatcccaactttaagaagaattatcgcttcagta 3596  
Db 2977 gctccctgcaaatatccagatggaagaatcccaactttaagaagaattatcgcttcagta 3036  
Qy 3597 atcaaatgtctacataatgataactatacttggtcttaagtatggtccagatcaaaagatt 3656  
Db 3037 atcaaatgtctacataatgataactatacttggtcttaagtatggtccagatcaaaagatt 3096  
Qy 3657 cgaatgatactgctcagatgagcagcaatgaaacatccatctcattcattcagtgta 3716  
Db 3097 cgaatgatactgctcagatgagcagcaatgaaacatccatctcattcattcagtgta 3156  
Qy 3717 catgtgttctactgttacaagaaaaaagaaggatataaatgtgcactgtacaactctatcca 3776  
Db 3157 catgtgttctactgttacaagaaaaaagaaggatataaatgtgcactgtacaactctatcca 3216  
Qy 3777 ggtgttttttgagacagtggaatgttaccatccaaagctggaattgtgcgggtgaaatgc 3836  
Db 3217 ggtgttttttgagacagtggaatgttaccatccaaagctggaattgtgcgggtgaaatgc 3276  
Qy 3837 cttaatgagcagcatctcagatgagtggaatgagacacatttctgtgtacagcaataag 3896  
Db 3277 cttaatgagcagcatctcagatgagtggaatgagacacatttctgtgtacagcaataag 3336  
Qy 3897 tgtcagaactccctctggaatggtctctggaacatlagatattcagatcaagcttca 3956  
Db 3337 tgtcagaactccctctggaatggtctctggaacatlagatattcagatcaagcttca 3396  
Qy 3957 ggaacatatagacagtggtgcacaaagcttggtccagacttattccggatcaatcaat 4016  
Db 3397 ggaacatatagacagtggtgcacaaagcttggtccagacttattccggatcaatcaat 3456  
Qy 4017 gctctgagcacaagaagaccttctctggaatcaagctgagatctgttgacacaaatgat 4076  
Db 3457 gctctgagcacaagaagaccttctctggaatcaagctgagatctgttgacacaaatgat 3516  
Qy 4077 attcaagcagcatcaagaaccagaggtgtcccgctcagaagttctccagcctctacactcag 4136  
Db 3517 attcaagcagcatcaagaaccagaggtgtcccgctcagaagttctccagcctctacactcag 3576  
Qy 4137 tttaatcatagtaagtctgtatggtggaagaagtgtgcagactatcgagggaattccact 4196  
Db 3577 tttaatcatagtaagtctgtatggtggaagaagtgtgcagactatcgagggaattccact 3636  
Qy 4197 ggaaccttaaatggtctcttctggaatgtgattcatctggaataaacaataatttt 4256  
Db 3637 ggaaccttaaatggtctcttctggaatgtgattcatctggaataaacaataatttt 3696  
Qy 4257 aacctccaattatgtctgcgaatacctggttgcacccaactcatatagcatltagcagc 4316  
Db 3697 aacctccaattatgtctgcgaatacctggttgcacccaactcatatagcatltagcagc 3756  
Qy 4317 acctctgcagatgagtgatggtgctgtgaatttaaatgttgacagatccatgtggaatg 4376

Db 3757 acctctgcagatgagtgatggtgctgtgaatttaaatgttgacagatccatgtggaatg 3816  
Qy 4377 gagagtaaaagcatalatcagatgcacagatctactgttcatctacttaccataatgttt 4436  
Db 3817 gagagtaaaagcatalatcagatgcacagatctactgttcatctacttaccataatgttt 3876  
Qy 4437 gccacctgtctctcttcaaaaagctcgaacttcaacctcaagggaggagtaatgtcctgaga 4496  
Db 3877 gccacctgtctctcttcaaaaagctcgaacttcaacctcaagggaggagtaatgtcctgaga 3936  
Qy 4497 cctcagggaataatccaaaagagtggtgtgaagtgtgcagcttccgaagacataaagtgc 4556  
Db 3937 cctcagggaataatccaaaagagtggtgtgaagtgtgcagcttccgaagacataaagtgc 3996  
Qy 4557 acaggagtaactactcaggaagataaatctcgtctaccagatgtatgtgaagagctc 4616  
Db 3997 acaggagtaactactcaggaagataaatctcgtctaccagatgtatgtgaagagctc 4056  
Qy 4617 ctcatctccagcagtcacaagatgtgcataagtggaactctcttcttccagaatgcaagta 4676  
Db 4057 ctcatctccagcagtcacaagatgtgcataagtggaactctcttcttccagaatgcaagta 4116  
Qy 4677 aaggttttcagggaataatcaagaactccttcaacactgtgtggaactctctagaccacag 4736  
Db 4117 aaggttttcagggaataatcaagaactccttcaacactgtgtggaactctctagaccacag 4176  
Qy 4737 ttactgactcgtctactcttccgaatltcaaccocagagltgggtgcacagatgtcccttgag 4796  
Db 4177 ttactgactcgtctactcttccgaatltcaaccocagagltgggtgcacagatgtcccttgag 4236  
Qy 4797 atggaagttctggtgtcgcaggacacaggaactctactga 4835  
Db 4237 atggaagttctggtgtcgcaggacacaggaactctactga 4275

RESULT 12  
AAN80446  
ID AAN80446 standard; DNA; 4275 BP.  
XX  
AC AAN80446;  
XX  
DT 10-OCT-1990 (first entry)  
XX  
DE Modified factor VIII:C sequence with the R740-E1649 deletion.  
XX  
KW Modified factor VIII:C; haemophilia; procoagulant;  
XX blood coagulation; RE deletion; ss.  
OS Homo sapiens.  
PN W08800831-A.  
XX  
PD 11-FEB-1988.  
XX  
PP 31-JUL-1987; 87WO-US01814.  
XX  
PR 01-AUG-1986; 86US-0893375.  
XX  
PA (BIOJ ) BIOGEN NV.  
XX  
PI Pasek MP;  
XX  
DR WPI: 1988-049866/07.  
XX  
XX P-PSDB: AAP80267.  
PT New DNA sequences encoding modified factor VIII:C - with deletion of DNA  
XX encoding maturation polypeptide, useful for high yield transformation.  
PS  
XX  
XX Claim 3; Page 44-45-46-47; 97pp; English.  
CC The entire sequence encoding the maturation polypeptide of  
XX factor VIII:C is deleted, i.e. Arg 740-Glu 1649.  
XX The full length Factor VIII:C cDNA has two changes with respect to the

CC published sequence (EPO application 1604577):  
CC CTC to CTA at Leu 242 and TTC to CTC change at amino acid residue 1880  
CC (Phe to Leu) : The product is produced in approx. 20 times higher  
CC yields than previous recombinant produced factor VIII:C and are more  
CC easily purified. The peptide is used for treating haemophilia A, both  
CC acute and prolonged bleeding.  
CC See also AAN80444 and AAN80447.  
CX  
SQ

Sequence 4275 BP; 1245 A; 940 C; 946 G; 1144 T; 0 other;

Query Match 52.5%; Score 4171.6; DB 9; Length 4275;  
Best Local Similarity 98.0%; Pred. No. 0;  
Matches 4268; Conservative 0; Mismatches 4; Indels 84; Gaps 1;

QY 477 gccacgaagaatactactcctcgtgctgagctggaactgtcatgtcagctatgtcaagaatgat 536  
DB 4 gccacgaagaatactactcctcgtgctgagctggaactgtcatgtcagctatgtcaagaatgat 63  
QY 537 ctccgttagctcctgtgtggaagcaagatlltcctcctagagtgaccaaatctttccattc 596  
DB 64 ctccgttagctcctgtgtggaagcaagatlltcctcctagagtgaccaaatctttccattc 123  
QY 597 aacacctcagtcgtgtacaaaagaactctgtttgtagaatcacggttcaccttttcaac 656  
DB 124 aacacctcagtcgtgtacaaaagaactctgtttgtagaatcacggttcaccttttcaac 183  
QY 657 atgcctagaccgaagccacccctgtagtggctgtgttagtcttaccatccaggctgaagtt 716  
DB 184 atgcctagaccgaagccacccctgtagtggctgtgttagtcttaccatccaggctgaagtt 243  
QY 717 tatgatacagtcgtcatctacacttaagaacatgtgcttccatccctgtcagctcatagct 776  
DB 244 tatgatacagtcgtcatctacacttaagaacatgtgcttccatccctgtcagctcatagct 303  
QY 777 gttgtgtatctcctctggaagactcttcggaagagctgaataatgtatgtatcagaaccatgcaa 836  
DB 304 gttgtgtatctcctctggaagactcttcggaagagctgaataatgtatgtatcagaaccatgcaa 363  
QY 837 agggagaagaagaatgtataaagcttcctcgtgtggaagccatacatatgtcttggaagctc 896  
DB 364 agggagaagaagaatgtataaagcttcctcgtgtggaagccatacatatgtcttggaagctc 423  
QY 897 ctgaagaagaatgttccaatgtgcctgtgacccactgtgccttacctactactatctctct 956  
DB 424 ctgaagaagaatgttccaatgtgcctgtgacccactgtgccttacctactactatctctct 483  
QY 957 catgtgacctgtgtaaaagacttgtaattcagggcctcatcttggaagccctactagatgtaga 1016  
DB 484 catgtgacctgtgtaaaagacttgtaattcagggcctcatcttggaagccctactagatgtaga 543  
QY 1017 gaaaggagctctggaaggaagaagaacacagacctgtgcaaaatttatactactctttgct 1076  
DB 544 gaaaggagctctggaaggaagaagaacacacacacctgtgcaaaatttatactactctttgct 603  
QY 1077 gtatttgaagaaggaaaagtctggacctcagaacaaagaactcctgtatgcagatag 1136  
DB 604 gtatttgaagaaggaaaagtctggacctcagaacaaagaactcctgtatgcagatag 663  
QY 1137 gatgtgtcatctgtctgggacctgtgcttaaaatgtcacacagtcnaatgtttatgttaaacagg 1196  
DB 664 gatgtgtcatctgtctgggacctgtgcttaaaatgtcacacagtcnaatgtttatgttaaacagg 723  
QY 1197 tctctgcaagtcctgtatgtatgacacaggaatacagtcctatctatgtgcatgtgagatg 1256  
DB 724 tctctcaacaagtcctgtatgtatgacacaggaatacagtcctatctatgtgcatgtgagatg 783  
QY 1257 ggcacacacctcctgaagtgcactcaatatctcctcgaaggtcacacatacttctgttgaagaa 1316  
DB 784 ggcacacacctcctgaagtgcactcaatatctcctcgaaggtcacacatacttctgttgaagaa 843  
QY 1317 catgcgcagagcttccttgggaatctgcacaaataacttctctactgtctaaacactcttg 1376

DB 844 catgcgcagagcttccttgggaatctgcacaaataacttctctactgtctaaacactcttg 903  
QY 1377 atgagacctggaagcttctctctgtttgttatatcttctcccaacaatgagatgctg 1436  
DB 904 atgagacctggaagcttctctctgtttgttatatcttctcccaacaatgagatgctg 963  
QY 1437 gaaacttatgtcaaaagttagaagctgtccagaaggaagcccaactcagatgaaatatat 1496  
DB 964 gaaacttatgtcaaaagttagaagctgtccagaaggaagcccaactcagatgaaatatat 1023  
QY 1497 gaagaagcggagaagactatgatgatgtcttaactgtatcttgaatgtatgtgtcaggttt 1556  
DB 1024 gaagaagcggagaagactatgatgatgtcttaactgtatcttgaatgtatgtgtcaggttt 1083  
QY 1557 gatgatgcaactctcctctcttattccaaatttgctgaagttgccaagaacatctctaa 1616  
DB 1084 gatgatgcaactctcctctcttattccaaatttgctgaagttgccaagaacatctctaa 1143  
QY 1617 acttggttacattacatgtctgtgaagagagactggaactatgtctcccttaagctctc 1676  
DB 1144 acttggttacattacatgtctgtgaagagagactggaactatgtctcccttaagctctc 1203  
QY 1677 gcccccgatgacagaagttataaagttaacataattgaaacatgtgcctcagcgagttgct 1736  
DB 1204 gcccccgatgacagaagttataaagttaacataattgaaacatgtgcctcagcgagttgct 1263  
QY 1737 aggaagtacaaaagaagtcctgttatgtgcatcaccaagttgaacctttaagaactgttaa 1796  
DB 1264 aggaagtacaaaagaagtcctgttatgtgcatcaccaagttgaacctttaagaactgttaa 1323  
QY 1797 gctatlcagcatgtaacaggaatcttggagacttaacttcttcttggtggaagcttggaagcca 1856  
DB 1324 gctatlcagcatgtaacaggaatcttggagacttaacttcttcttggtggaagcttggaagcca 1383  
QY 1857 ctgtgttatataatttaagaatccaagcaagcaacataatacatctacccctacaggaatc 1916  
DB 1384 ctgtgttatataatttaagaatccaagcaagcaacataatacatctacccctacaggaatc 1443  
QY 1917 actgatgtccgtctcttctgtattcacaaggagttaccaaaaggtgtaaacatttgaagat 1976  
DB 1444 actgatgtccgtctcttctgtattcacaaggagttaccaaaaggtgtgtaaacatttgaagat 1503  
QY 1977 ttcccaattctgcagaggaataattcaaalataatgtgaacatgtcgttagaagtgtg 2036  
DB 1504 ttcccaattctgcagaggaataattcaaalataatgtgaacatgtcgttagaagtgtg 1563  
QY 2037 ccaactaaatcagatcctcgtgtcctgtgacccgctattactcttagtttgcgttaatatgag 2096  
DB 1564 ccaactaaatcagatcctcgtgtcctgtgacccgctattactcttagtttgcgttaatatgag 1623  
QY 2097 agagatctaagctttagaagcatattggccctctcctcatctgtcaaaagaatctgtagat 2156  
DB 1624 agagatctaagctttagaagcatattggccctctcctcatctgtcaaaagaatctgtagat 1683  
QY 2157 caaagaggaacacagataatgtlccagacaagagagatgtcaccctgttttcgtaattgat 2216  
DB 1684 caaagaggaacacagataatgtlccagacaagagagatgtcaccctgttttcgtaattgat 1743  
QY 2217 gagaacccgaagctgtgtacctcacaagagaatatatacaagcttctccccaatccagctgga 2276  
DB 1744 gagaacccgaagctgtgtacctcacaagagaatatatacaagcttctccccaatccagctgga 1803  
QY 2277 gtgcaagctttagaagttcagaagttccaggccccaactcatgtgacacagatcaatagtgctat 2336  
DB 1804 gtgcaagctttagaagttcagaagttccaggccccaactcatgtgacacagatcaatagtgctat 1863  
QY 2337 gtttttgatagttgtgcagttgtcagttgtttgtgatagtgaggtgacatctgatactcta 2396  
DB 1864 gtttttgatagttgtgcagttgtcagttgtttgtgatagtgaggtgacatctgatactcta 1923  
QY 2397 agcatgtgagacagagactgactcctctctgtctctctctcgtgatataccttcaaacac 2456  
DB 1924 agcatgtgagacagagactgactcctctctgtctctctctcgtgatataccttcaaacac 1983

|    |      |  |      |
|----|------|--|------|
| QY | 2457 | aaaaatgctctatgtaaagacaacaactccaaccttatcccatcttccaaggagaacatcgcttctaatg | 2516 |
| Db | 1984 | aaaaatgctctatgtaaagacaacaactccaaccttatcccatcttccaaggagaacatcgcttctaatg | 2043 |
| QY | 2517 | tgcagtgaagaaacccaagcttcctatgactcttgggtgtgcacaacatcagaacttcgcgaacaga    | 2576 |
| Db | 2044 | tgcagtgaagaaacccaagcttcctatgactcttgggtgtgcacaacatcagaacttcgcgaacaga    | 2103 |
| QY | 2577 | ggcagtagaccgccttactgaagagttctctagtgtgtgacaagaacacttgtgtattatcagag      | 2636 |
| Db | 2104 | ggcagtagaccgccttactgaagagttctctagtgtgtgacaagaacacttgtgtattatcagag      | 2163 |
| QY | 2637 | ggcagtagtagaagaattttcagaatacttctcagtagtaaaaaacaattgcacatggaacacaga     | 2696 |
| Db | 2164 | ggcagtagtagaagaattttcagaatacttctcagtagtaaaaaacaattgcacatggaacacaga     | 2223 |
| QY | 2697 | agcttctccagaagatttcaagacacccctagactagtgaaaaagcaatttaattgcacccca        | 2756 |
| Db | 2224 | -----  | 2223 |
| QY | 2757 | ccaagcttctgaaagcccaatcaacgggaaataactccgtactactcttccaagtccaagag         | 2816 |
| Db | 2224 | -----gaataactcgtactactcttccaagtccaagag                                 | 2259 |
| QY | 2817 | gaatatgactatgtatgataacatcatcactcgtgaaatggaagaggaagatttgcacattat        | 2876 |
| Db | 2260 | gaatatgactatgtatgataacatcatcactcgtgaaatggaagaggaagatttgcacattat        | 2319 |
| QY | 2877 | gatgagatgataaaatcagagcccccgagcttccaagaagaacaacagacattattat             | 2936 |
| Db | 2320 | gatgagatgataaaatcagagcccccgagcttccaagaagaacaacagacattattat             | 2379 |
| QY | 2937 | gctgcacgtggaagaggtctctgggtgtatbtggatgtagatgcctcccaacatgtttcctaagaac    | 2996 |
| Db | 2380 | gctgcacgtggaagaggtctctgggtgtatbtggatgtagatgcctcccaacatgtttcctaagaac    | 2439 |
| QY | 2997 | agggtctcaagttgacagtgctccctcagtttcaagaagaattgttttccagagaattaccgat       | 3056 |
| Db | 2440 | agggtctcaagttgacagtgctccctcagtttcaagaagaattgttttccagagaattaccgat       | 2499 |
| QY | 3057 | ggctcccttactcagcccttataccgtggtgagaacttaataacatttggagatctctggg          | 3116 |
| Db | 2500 | ggctcccttactcagcccttataccgtggtgagaacttaataacatttggagatctctctggg        | 2559 |
| QY | 3117 | ccatataatagcagagagttgtagaataatcatcatgtttaacttccaagaacatcagacct         | 3176 |
| Db | 2560 | ccatataatagcagagagttgtagaataatcatcatgtttaacttccaagaacatcagacct         | 2619 |
| QY | 3177 | gcttccctatctcttcaattcttaagccttatcttctatgtagaagaagatcagaaggcaaggagca    | 3236 |
| Db | 2620 | gcttccctatctcttcaattcttaagccttatcttctatgtagaagaagatcagaaggcaaggagca    | 2679 |
| QY | 3237 | ggaacctagaanaaaactttgtcagagcccaatggaacccaacttbaacttcttggaaagtcaa       | 3296 |
| Db | 2680 | ggaacctagaanaaaactttgtcagagcccaatggaacccaacttbaacttcttggaaagtcaa       | 2739 |
| QY | 3297 | catcataatggtccaccaactaaagatgttgcactgtgaagccttgggcttatctctgat           | 3356 |
| Db | 2740 | catcataatggtccaccaactaaagatgttgcactgtgaagccttgggcttatctctgat           | 2799 |
| QY | 3357 | gttgcacctggaanaaaatgtgtgcactgaagccttgaatggaccccttctgtgtcgtcgaacct      | 3416 |
| Db | 2800 | gttgcacctggaanaaaatgtgtgcactgaagccttgaatggaccccttctgtgtcgtcgaacct      | 2859 |
| QY | 3417 | aacacacttaaaccttctcactatgtaggaagacaagtacacatagcagaattctgccttcttc       | 3476 |
| Db | 2860 | aacacactgaaaccttctcactatgtaggaagacaagtacacatagcagaattctgtcttcttc       | 2919 |
| QY | 3477 | accacatcttggatgtagaacccaagaagctgtgtacttcaactgtgaaaaatgtgaagaacatcgag   | 3536 |
| Db | 2920 | accacatcttggatgtagaacccaagaagctgtgtacttcaactgtgaaaaatgtgaagaacatcgag   | 2979 |

|    |      |   |      |
|----|------|---|------|
| OY | 3537 | gcccctgcgaatatalccagatctggaagatcccaacttttaagaagatlatcgtcttcatactga  | 3556 |
| Db | 2980 | gtcccttcgcaatatccagatctggaagatcccaacttttaagaagatlatcgtcttcatactga   | 3039 |
| OY | 3537 | atcaatgctgaataaagatgaataacactggttgatgaatgctccagatccaagaagt          | 3656 |
| Db | 3040 | atcaatgctgaataaagatgaataacactggttgatgaatgctccagatccaagaagt          | 3099 |
| OY | 3657 | cgatggtatctcgtccagatcgaggcagcaaatgaaataatccattcatcttcaatcttcagtga   | 3716 |
| Db | 3100 | cgatggtatctcgtccagatcgaggcagcaaatgaaataatccattcatcttcaatcttcagtga   | 3159 |
| OY | 3717 | catgtgttcaactgtatcgaaaaaagagatlaaaatggaactgtacatctcatcca            | 3776 |
| Db | 3160 | catgtgttcaactgtatcgaaaaaagagatlaaaatggaactgtacatctcatcca            | 3219 |
| OY | 3777 | gggttttttgagaagatggaatgttatccatccaaagctggaatttggcgggtggaatgc        | 3856 |
| Db | 3220 | gggttttttgagaagatggaatgttatccatccaaagctggaatttggcgggtggaatgc        | 3279 |
| OY | 3837 | ctgatctggcgagatctacatactcgttgatgagacaacttttctgtgtacagcaataag        | 3896 |
| Db | 3280 | ctgatctggcgagatctacatactcgttgatgagacaacttttctgtgtacagcaataag        | 3339 |
| OY | 3897 | tgatcagaatcccccctgbygaatggtctctctggacacatgagaatlttcagattacagcttca   | 3956 |
| Db | 3340 | tgatcagaatcccccctgbygaatggtctctctggacacatgagaatlttcagattacagcttca   | 3399 |
| OY | 3957 | ggacaatatatgacagatggggccccaagctgycgaagcttcattatctccggtatcaatcaat    | 4016 |
| Db | 3400 | ggacaatatatgacagatggggccccaagctgycgaagcttcattatctccggtatcaatcaat    | 3459 |
| OY | 4017 | gcccggagacccaagggagcccttctcttgatcaaaagtgtgaactgtgttgacaacatgatt     | 4076 |
| Db | 3460 | gcccggagacccaagggagcccttctcttgatcaaaagtgtgaactgtgtgtgacaacatgatt    | 3519 |
| OY | 4077 | attcacgacatcaagaccccaagggtggcccgatcagaagttctccagcctctacatctctcaag   | 4136 |
| Db | 3520 | attcacgacatcaagaccccaagggtggcccgatcagaagttctccagcctctacatctctcaag   | 3579 |
| OY | 4137 | tttatcatcatgtatagttctgtatgggaagaagtggcagacttatccggagaatttcaact      | 4196 |
| Db | 3560 | tttatcatcatgtatagttctgtatgggaagaagtggcagacttatccggagaatttcaact      | 3639 |
| OY | 4197 | ggaacccaatgagctcctctcttgcaaatgttgatattcatctcggagataaacaacaatatcttt  | 4256 |
| Db | 3640 | ggaacccaatgagctcctctcttgcaaatgttgatattcatctcggagataaacaacaatatcttt  | 3699 |
| OY | 4257 | aaacccaatatttgctcgtacatacatccgttttgacccaactcattatagatcttcagac       | 4316 |
| Db | 3700 | aaacccaatatttgctcgtacatacatccgttttgacccaactcattatagatcttcagac       | 3759 |
| OY | 4317 | actcttcgcatagtatgttgatgggctgtgtgtttaaatagtgtgagaatgccaatlggnaatg    | 4376 |
| Db | 3760 | actcttcgcatagtatgttgatgggctgtgtgtttaaatagtgtgagaatgccaatlggnaatg    | 3819 |
| OY | 4377 | gagagtaaaacaatatcatagaatgcagatgtaactggtcttacccttaccataatatgctt      | 4436 |
| Db | 3820 | gagagtaaaacaatatcatagaatgtaactggtcttacccttaccataatatgctt            | 3879 |
| OY | 4437 | ggcaacctggtctcccttccaagaagctcgtactccaactccaagggagagatgaatgycctbgaga | 4496 |
| Db | 3880 | ggcaacctggtctcccttccaagaagctcgtactccaactccaagggagagatgaatgycctbgaga | 3939 |
| OY | 4497 | cctcaggtgtaataatccaagaagatggtctgcaagtggacttccagaaagacaatgaagaatc    | 4556 |
| Db | 3940 | cctcaggtgtaataatccaagaagatggtctgcaagtggacttccagaaagacaatgaagaatc    | 3999 |
| OY | 4557 | acaagagtaactactcctcagggagatgaataatctcgtcttacaagaaatgatgtgaagaagctc  | 4616 |
| Db | 4000 | acaagagtaactactcctcagggagatgaataatctcgtcttacaagaaatgatgtgaagaagctc  | 4059 |
| OY | 4617 | ctcatctccacagagatcaagaatggccatctcagtggaactctctttttccagaatggaagaatga | 4676 |



|||||  
Db 4060 ctaatccacagcagtaagatgagcagatgagctctcttttccagaatgccaagtta 4119  
QY 4677 aagattttcaggaatacaagactccttcacacctgtgtgtaactctagaccacgc 4736  
Db 4120 aagtttttcagggaatacaagactccttcacacctgtgtgtaactcttagaccacgc 4179  
QY 4737 ttactgactcgtactactcgtatctacacccagagttgggtgacacagattgccttagg 4796  
Db 4180 ttactgactcgtactactcgtatctacacccagagttgggtgacacagattgccttagg 4239  
QY 4797 atggaagttctgggctgcgaagcagacagacctcac 4832  
Db 4240 atggaagttctgggctgcgaagcagacagacctcac 4275  
  
RESULT 13  
AAN80447  
ID AAN80447 standard; DNA; 4272 BP.  
XX  
AC AAN80447;  
XX  
DT 10-OCT-1990 (first entry)  
XX  
DE Modified factor VIII:C sequence with the R740-D1658 deletion.  
XX  
KM Modified factor VIII:C; haemophilia; procoagulant;  
KM blood coagulation; RD deletion; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO8800831-A.  
XX  
PD 11-FEB-1988.  
XX  
PF 31-JUL-1987; 87WO-US01814.  
XX  
PR 01-AUG-1986; 86US-0893375.  
XX  
PA (BIOJ ) BIOGEN NV.  
XX  
PI Pasek MP;  
XX  
DR WPI: 1988-049866/07.  
XX  
PT P-PSDB; AAP80268.  
XX  
PS New DNA sequences encoding modified factor VIII:C - with deletion of DNA  
encoding maturation polypeptide, useful for high yield transformation.  
XX  
PS Claim 3; Page 47-48-49-50; 97pp; English.  
XX  
CC The RD deletion removes the DNA from Ser 741 to Ser 1657. A major part  
of the sequence encoding the maturation polypeptide of  
CC factor VIII:C is deleted, i.e. Gln 744 - Asp 1563.  
CC The full length Factor VIII:C cDNA has two changes with respect to the  
CC published sequence (EPO application 160457):  
CC CNG to CNA at Leu 242 and TTC to CTC change at amino acid residue 1850  
CC (Phe to Leu). The product is produced in approx. 20 times higher  
CC yields than previous recombinant produced factor VIII:C and are more  
CC easily purified. The peptide is used for treating haemophilia A, both  
CC acute and prolonged bleeding.  
CC See also AAN80444 and AAN80446.  
XX  
SQ Sequence 4272 BP; 1243 A; 941 C; 944 G; 1144 T; 0 other;

Query Match 52.5%; Score 4170; DB 9; Length 4272;  
Best Local Similarity 98.0%; Pred. No. 0;  
Matches 4267; Conservative 0; Mismatches 5; Indels 84; Gaps 1;

QY 477 gccacagaatactactcgtggtgcagtggaactgtcatggaactatgcaaatgcat 536  
|||||  
Db 1 gccacagaatactactcgtggtgcagtggaactgtcatggaactatgcaaatgcat 60

QY 537 ctggtgagctgctgtggagcagaagatttctctctagagtgcacaaatctttcattc 596  
Db 61 ctggtgagctgctgtggagcagaagatttctctctagagtgcacaaatctttcattc 120  
QY 597 aacacccagtcgtgtacaaagaacctgtgtttaaatccaagttcacctttcaac 656  
Db 121 aacacccagtcgtgtacaaagaacctgtgtttaaatccaagttcacctttcaac 180  
QY 657 atcgtaagccaagccaacctgtgattggtctgtcttaagttccaccatccagctgaagt 716  
Db 181 atcgtaagccaagccaacctgtgattggtctgtcttaagttccaccatccagctgaagt 240  
QY 717 tatgatacagtggtcattacacttaagaactgtgtccatctctgtcagcttcatgct 776  
Db 241 tatgatacagtggtcattacacttaagaactgtgtccatctctgtcagcttcatgct 300  
QY 777 gtgtgtatccctactgaaagctcttgaggagctgaatagtatgatacagcagtcacaa 836  
Db 301 gtgtgtatccctactgaaagctcttgaggagctgaatagtatgatacagcagtcacaa 360  
QY 837 agggagaagaagaatgaataagttctccctggtggaagccatacatatgtctggcaggtc 896  
Db 361 agggagaagaagaatgaataagttctccctggtggaagccatacatatgtctggcaggtc 420  
QY 897 ctgaaagagaatgtgccaaatggtcctgcagccactgtgcttactactatcttcttct 956  
Db 421 ctgaaagagaatgtgccaaatggtcctgcagccactgtgcttactactatcttcttct 480  
QY 957 catgtgacctgtgtaaaagactgaattcagagctcactatgagccactaagtatgata 1016  
Db 481 catgtgacctgtgtaaaagactgaattcagagctcactatgagccactaagtatgata 540  
QY 1017 gaaggagctctgccaaggaagaacacagacctgtgcaacaatttactacttttgc 1076  
Db 541 gaaggagctctgccaaggaagaacacacacctgtgcaacaatttactacttttgc 600  
QY 1077 gtatttgatgaaggaaagtgtgactctcagaacaaagaactccttgatgcaagatagg 1136  
Db 601 gtatttgatgaaggaaagtgtgactctcagaacaaagaactccttgatgcaagatagg 660  
QY 1137 gatgtcatctgtctggcgcctgacctaaatgacacacagtcacatgtgtatgatacagg 1196  
Db 661 gatgtcatctgtctggcgcctgacctaaatgacacacagtcacatgtgtatgatacagg 720  
QY 1197 tctctgcaagttctgattgtagtcacaggaatcagtcctatggtcagtgtgatagt 1256  
Db 721 tctctgcaagttctgattgtagtcacaggaatcagtcctatggtcagtgtgatagt 780  
QY 1257 ggcacacactcctgaagtgcaactaatttctctgaaagtgcaacaaatttctgtgagac 1316  
Db 781 ggcacacactcctgaagtgcaactaatttctctgaaagtgcaacaaatttctgtgagac 840  
QY 1317 catcgcaagcgctctctgggaatctgcacataacttctcttactgtctcaaacctcttg 1376  
Db 841 catcgcaagcgctctctgggaatctgcacataacttctcttactgtctcaaacctcttg 900  
QY 1377 atggaactgtggaacttctactgttttgcataatctcttccacacaatgatagtcagat 1436  
Db 901 atggaactgtggaacttctactgttttgcataatctcttccacacaatgatagtcagat 960  
QY 1437 gaagcttatgtcaaaagtagcagctgtccagaggaaccccaactcagaaatgaanaaat 1496  
Db 961 gaagcttatgtcaaaagtagcagctgtccagaggaaccccaactcagaaatgaanaaat 1020  
QY 1497 gaagaagcgaagaactatgatacttactcagatctgaatgtgatagtgtcaggttc 1556  
Db 1021 gaagaagcgaagaactatgatacttactcagatctgaatgtgatagtgtcaggttc 1080  
QY 1557 gatgatgaacaactctctcttctatccaaatctgcgtcagttgcgaagaacatcctaa 1616  
Db 1081 gatgatgaacaactctctcttctatccaaatctgcgtcagttgcgaagaacatcctaa 1140



|    |      |   |      |
|----|------|---|------|
| QY | 1617 | acttgaggtacattacatctgtcctcggaagagagacttggagactatgtcccttagtcttc      | 1676 |
| Db | 1141 | acttgggtacattacatctgtcctcggaagagagagacttggagactatgtcccttagtcttc     | 1200 |
| QY | 1677 | gcccccatgacaaagtataaagaatcaatttgaacaaatggccctcagcgatggt             | 1736 |
| Db | 1201 | gcccccatgacaaagtataaagaatcaatttgaacaaatggccctcagcgatggt             | 1260 |
| QY | 1737 | aggaagtacaaaagaftccgaatttatgtgcatacacagaatgaaccttaagaactcgttga      | 1796 |
| Db | 1261 | aggaagtacaaaagaftccgaatttatgtgcatacacagaatgaaccttaagaactcgttga      | 1320 |
| QY | 1797 | gctattcaagcatgataccaagaactcttggaccttacttatttgggaagttggagacaca       | 1856 |
| Db | 1321 | gctattcaagcatgataccaagaactcttggaccttacttatttgggaagttggagacaca       | 1380 |
| QY | 1857 | ctgtgtgattatatttaagaatccaagcaagcgagacatatcaactaccctccaggaatc        | 1918 |
| Db | 1381 | ctgtgtgattatatttaagaatccaagcaagcgagacatatcaactaccctccaggaatc        | 1440 |
| QY | 1917 | actgtatgtccgctcttctgtattccaaggagatataccaagaagtgttaaaacatttgaagat    | 1976 |
| Db | 1441 | actgtatgtccgctcttctgtattccaaggagatataccaagaagtgttaaaacatttgaagat    | 1500 |
| QY | 1977 | tttccaattcttcgcagagagaaatatccaatatataatgtgacaagtgactgttgaagatg      | 2036 |
| Db | 1501 | tttccaattcttcgcagagagaaatatccaatatataatgtgacaagtgactgttgaagatg      | 1560 |
| QY | 2037 | ccaactaatatagaatccctcggtgcctgcagccgcctatctcttagtttcgttaaatgtgag     | 2096 |
| Db | 1561 | ccaactaatatagaatccctcggtgcctgcagccgcctatctcttagtttcgttaaatgtgag     | 1620 |
| QY | 2097 | agaagatcctaagcactcaatcttgcgcctctccatcaltctgtacaagaatctgttagat       | 2156 |
| Db | 1621 | agaagatcctaagcactcaatcttgcgcctctccatcaltctgtacaagaatctgttagat       | 1680 |
| QY | 2157 | caaaagagaaaaccaagataatgtcagaacaagaagatgtacatccgttttctgtatttgat      | 2216 |
| Db | 1681 | caaaagagaaaaccaagataatgtcagaacaagaagatgtacatccgttttctgtatttgat      | 1740 |
| QY | 2217 | gagaacccgaagctgtgtacccccaagagaaatatcaaacgcttctccccaacccagctgtga     | 2276 |
| Db | 1741 | gagaacccgaagctgtgtacccccaagagaaatatcaaacgcttctccccaacccagctgtga     | 1800 |
| QY | 2277 | gttcgagctcttgagataccaagattccaagcctccaacatcaltgcagaagatcaatgtgcat    | 2336 |
| Db | 1801 | gttcgagctcttgagataccaagattccaagcctccaacatcaltgcagaagatcaatgtgcat    | 1866 |
| QY | 2337 | gtttctgtatgtttgcagagctgtgcaggtttgttcttcgataaggtgtgcatactgttaccatca  | 2396 |
| Db | 1861 | gtttctgtatgtttgcagagctgtgcaggtttgttcttcgataaggtgtgcatactgttaccatca  | 1920 |
| QY | 2397 | agcatctgagacacagcactgacttccctctcgtctcttcctctctgtgataactccaacac      | 2456 |
| Db | 1921 | agcatctgagacacagcactgacttccctctcgtctcttcctctctgtgataactccaacac      | 1980 |
| QY | 2457 | aaaatgttctatgagagacacactcaacctatcccatctccaaggaacgtcttcatgt          | 2516 |
| Db | 1981 | aaaatgttctatgagagacacactcaacctatcccatctccaaggaacgtcttcatgt          | 2040 |
| QY | 2517 | tcgatagtgaaaaccaggtctatggaatctctgggtgtgccaaaccacccagacttcctggaacaga | 2576 |
| Db | 2041 | tcgatagtgaaaaccaggtctatggaatctctgggtgtgccaaaccacccagacttcctggaacaga | 2100 |
| QY | 2577 | ggcatgacgcgcttactcgaaggttttctagtgtgtacaagaacactggtgtattatacgag      | 2636 |
| Db | 2101 | ggcatgacgcgcttactcgaaggttttctagtgtgtacaagaacactggtgtattatacgag      | 2160 |
| QY | 2637 | gacagctatgaaagatatcttcagcaactactgtctgggttaaaaacaaatgcatgtgaaccaga   | 2696 |
| Db | 2161 | gacagctatgaaagatatcttcagcaactactgtctgggttaaaaacaaatgcatgtgaaccaga   | 2220 |
| QY | 2697 | agcttctcccaagaattccaagacacctctagcctctgtgcagaagaatttaatgtcacccca     | 2756 |

|    |      |  |      |
|----|------|--|------|
| Db | 2221 | -----  | 2220 |
| Qy | 2757 | coagcttgaagacgcatacagcgaataataactcgttactactcttcaagtcagatcaagag<br> | 2816 |
| Db | 2221 | -----  | 2256 |
| Qy | 2817 | gaaatgtacatactgatbaccatactcaagttgaatgaagaggaagatttgcacattat        | 2876 |
| Db | 2257 | gaaatgtacatactgatbaccatactcaagttgaatgaagaggaagatttgcacattat        | 2316 |
| Qy | 2877 | gataggaatgaataatcagagcccccagagcttccaagaagaacccgacattttat           | 2936 |
| Db | 2317 | gataggaatgaataatcagagcccccagagcttccaagaagaacccgacattttat           | 2376 |
| Qy | 2937 | gctgcagagcagagagcctcgtggaattatctggaatgtagcagctcccaatcttcaagaac     | 2996 |
| Db | 2377 | gctgcagagcagagagcctcgtggaatgtagcagctcccaatcttcaagaac               | 2436 |
| Qy | 2997 | agggtcagagtgagcagtgctccctcagttcaagaagattgttttccaagaaattacgtat      | 3056 |
| Db | 2437 | agggtcagagtgagcagtgctccctcagttcaagaagattgttttccaagaaattacgtat      | 2496 |
| Qy | 3057 | ggctccttactcagcccttaacccgtagaggaactaaatgaacatttggagactccctggg      | 3116 |
| Db | 2497 | ggctccttactcagcccttaacccgtagaggaactaaatgaacatttggagactccctggg      | 2556 |
| Qy | 3117 | ccataataagagcagagagttgaagataatataatgttaacttccaagaatcagccct         | 3176 |
| Db | 2557 | ccataataagagcagagagttgaagataatataatgttaacttccaagaatcagccct         | 2616 |
| Qy | 3177 | cgctccctatctccctcatctcagccctatttcttaatgtaggaagatacagagcagagca      | 3236 |
| Db | 2617 | cgctccctatctccctcatctcagccctatttcttaatgtaggaagatacagagcagagca      | 2676 |
| Qy | 3237 | gaacctgagaaaaaactttgttcagagccctaagaaaccaaacttactcttggagaagtgc      | 3296 |
| Db | 2677 | gaacctgagaaaaaactttgttcagagccctaagaaaccaaacttactcttggagaagtgc      | 2736 |
| Qy | 3297 | catcataagcagccactcaagaatgtagttgtaactgcaaaagccctggagcttattccgtat    | 3356 |
| Db | 2737 | catcataagcagccactcaagaatgtagttgtaactgcaaaagccctggagcttattccgtat    | 2796 |
| Qy | 3357 | gttgaacctgysaaaaagatgtgtgcactcagccgtatgtgagccctctgtgtctgcacact     | 3416 |
| Db | 2797 | gttgaacctgysaaaaagatgtgtgcactcagccgtatgtgagccctctgtgtctgcacact     | 2856 |
| Qy | 3417 | aacacactgaaacctgtctcatcttggagacaaagttaacagtcacaggaatttgcctgttttc   | 3476 |
| Db | 2857 | aacacactgaaacctgtctcatcttggagacaaagttaacagtcacaggaatttgcctgttttc   | 2916 |
| Qy | 3477 | aacacacttltgatgtgaccaaagctgtgtlaactaacgcgaataatagtgaagaacatgcag    | 3536 |
| Db | 2917 | aacacacttltgatgtgaccaaagctgtgtlaactaacgcgaataatagtgaagaacatgcag    | 2976 |
| Qy | 3537 | gtccctctgcaaatcaccagatgtgaagatccaccatttaagaggaatbatactgttccatgca   | 3596 |
| Db | 2977 | gtccctctgcaaatcaccagatgtgaagatccaccatttaagaggaatbatactgttccatgca   | 3036 |
| Qy | 3597 | atcaaatgtcataataatgtgatacactacccctgtgcttgaatgtgtccaggaatcaaaagga   | 3656 |
| Db | 3037 | atcaaatgtcataataatgtgatacactacccctgtgcttgaatgtgtccaggaatcaaaagga   | 3096 |
| Qy | 3657 | cgaatgttatctgtcgcacagatgaggcaggaatgaagaaatccatcttcatcttgaatttgg    | 3716 |
| Db | 3097 | cgaatgttatctgtcgcacagatgaggcaggaatgaagaaatccatcttcatcttgaatttgg    | 3156 |
| Qy | 3717 | catgtgttacctgtcagcaaaaaaagagagatataaaatgscctgtacacatctccatca       | 3776 |
| Db | 3157 | catgtgttacctgtcagcaaaaaaagagagatataaaatgscctgtacacatctccatca       | 3216 |
| Qy | 3777 | gggtgttcttgcagacatgtgaatgttaccatcccaagcctgtgaatttbgcgggtgaaatgc    | 3836 |

```
Db 3217 gggtttttgagacagtgaatgttacatccaagctggaatttggcgggtgaaatgc 3276
Oy 3837 cttaattggcagacatcctacgtcgggatgagccactttttcgggtgcagcaataag 3896
Db 3277 cttaattggcagacatcctacgtcgggatgagccactttttcgggtgcagcaataag 3336
Oy 3897 tgltagactccccggggaatggtctctgcagacataagagatttccagattacagcttca 3956
Db 3337 tgltagactccccggggaatggtctctgcagacataagagatttccagattacagcttca 3396
Oy 3957 ggcacataatgacagtgggcccacaagctggccagacttcaattcccgatcaatcaat 4016
Db 3397 ggcacataatgacagtgggcccacaagctggccagacttcaattcccgatcaatcaat 3456
Oy 4017 ggcctggagccccaagggccctttcttgatcaaggltgatactgtttggccaatgatc 4076
Db 3457 ggcctggagccccaagggccctttcttgatcaaggltgatactgtttggccaatgatc 3516
Oy 4077 attcagcgatcaagaccaggggtgcgcgtcagaagtctccacagcctctacatctcag 4136
Db 3517 attcagcgatcaagaccaggggtgcgcgtcagaagtctccacagcctctacatctcag 3576
Oy 4137 ttatcatcaltgatagttcttgatgggaagaagtggcagacttaccaggaattccact 4196
Db 3577 ttatcatcaltgatagttcttgatgggaagaagtggcagacttaccaggaattccact 3636
Oy 4197 ggaaccttaatggtcttcttgatgggaagtgtgattcattctgggataaaacaataatttc 4256
Db 3637 ggaaccttaatggtcttcttgatgggaagtgtgattcattctgggataaaacaataatttc 3696
Oy 4257 aacctccaattatgtctgcatacatcgtttgcaccaactcaattatagatcgcagc 4316
Db 3697 aacctccaattatgtctgcatacatcgtttgcaccaactcaattatagatcgcagc 3756
Oy 4317 actcttcgcatgagtgatgggtcgtgtgatttaaatgttgcagcatgcatgtggaaatg 4376
Db 3757 actcttcgcatgagtgatgggtcgtgtgatttaaatgttgcagcatgcatgtggaaatg 3816
Oy 4377 gtaggttaagaacatatacagatgcagattactgctcattcattaccataatttc 4436
Db 3817 gtaggttaagaacatatacagatgcagattactgctcattcattaccataatttc 3876
Oy 4437 gccacgtgctctcttcaaaaagctcgcacttcaacctccaagggagtaattgcctggaga 4496
Db 3877 gccacgtgctctcttcaaaaagctcgcacttcaacctccaagggagtaattgcctggaga 3936
Oy 4497 cctcaggtgataatccaagaagtggtcgtcgaagtggacttccagaagaacatgaaatgc 4556
Db 3937 cctcaggtgataatccaagaagtggtcgtcgaagtggacttccagaagaacatgaaatgc 3996
Oy 4557 acaggaatgactactcagggagtaaaatctcgttccagcagcatgtatgtgaaggagttc 4616
Db 3997 acaggaatgactactcagggagtaaaatctcgttccagcagcatgtatgtgaaggagttc 4056
Oy 4617 ctcattccacagcagtcagaatggtccatcaagtggaactctcttttcagaatggaagtga 4676
Db 4057 ctcattccacagcagtcagaatggtccatcaagtggaactctcttttcagaatggaagtga 4116
Oy 4677 aaggttttcaagggaataccaagactcctcacaactgtgtgtaactctcagaaccacg 4736
Db 4117 aaggttttcaagggaataccaagactcctcacaactgtgtgtaactctcagaaccacg 4176
Oy 4737 ttactagctgcgtacacttgcaattcaccccgagatgggttgacccagatgcccctgagg 4796
Db 4177 ttactagctgcgtacacttgcaattcaccccgagatgggttgacccagatgcccctgagg 4236
Oy 4797 atggaagttctggcgtcgcagagcaggaacctctac 4832
Db 4237 atggaagttctggcgtcgcagagcaggaacctctac 4272
```

RESULT 14  
AAx82258

```
ID AAX82258 standard; cDNA; 4373 BP.
XX AC
XX AAX82258;
XX DT 18-AUG-1999 (first entry)
XX DE Beta-domain deleted Factor VIII protein encoding cDNA.
XX KW Factor VIII protein; gene modification; gene therapy; clinical disorder;
XX KW splicing pattern; RNA processing; gene regulation; beta-domain; human;
XX OS Homo sapiens.
XX PN M09929848-A1.
XX PD 17-JUN-1999.
XX PF 25-NOV-1998; 98MO-US25354.
XX PR 16-JAN-1998; 98US-0071596.
XX PR 05-DEC-1997; 97US-0067614.
XX PA (IMMU-) IMMUNE RESPONSE CORP.
XX PI Bldingmaier S, Gonzales JEN, Ill CR, Yang CO;
XX WPI; 1999-385602/32.
XX DR P-PDB; AAY21675.
XX PT Genes and vectors exhibiting increased expression and novel splicing
XX PT patterns, useful for expression of, e.g. beta-domain deleted factor
XX PT VIII
XX PS Claim 9; Page 72-78; 123pp; English.
XX PS CC The invention describes novel genes and vectors exhibiting increased
XX CC expression and novel splicing patterns. It provides a gene encoding a
XX CC Factor VIII protein, that comprises one or more consensus or near
XX CC consensus splice sites which have been corrected to increase expression.
XX CC The method, DNA sequences and expression vectors can be used to increase
XX CC the expression of a gene, especially a Factor VIII gene. Genes containing
XX CC modified 5' and/or 3' untranslated regions have optimized expression
XX CC levels and tissue-specific expression. The methods are used for
XX CC identification and correction of consensus splice sites; addition of
XX CC introns; optimization of 5' and 3' untranslated regions and increase in
XX CC cytoplasmic RNA accumulation. Hence the DNAs are useful in gene therapy
XX CC to treat a clinical disorder, to study RNA processing and/or gene
XX CC regulation. The present sequence represents a cDNA encoding a beta-domain
XX CC deleted factor VIII protein.
XX SQ Sequence 4373 BP; 1278 A; 1011 C; 938 G; 1146 T; 0 other;

Query Match 52.4%; Score 4161; DB 20; Length 4373;
Best Local Similarity 96.8%; Pred. No. 0;
Matches 4273; Conservative 0; Mismatches 100; Indels 42; Gaps 1;
```

|    |      |  |      |
|----|------|--|------|
| Oy | 660  | gctaaagccaaagccacccttgatcgggtctctctagctccatccaaagctgaagttat        | 713  |
| Db | 241  | gctaaagccaaagccacccttgatcgggtctctctagctccatccaaagctgaagttat        | 300  |
| Oy | 720  | gatacagtgctaatcaacttaagaacatggtccctccatccctgcagttcttcagctgt        | 779  |
| Db | 301  | gatacagtgctaatcaacttaagaacatggtccctccatccctgcagttcttcagctgt        | 360  |
| Oy | 780  | ggtgtatctcacttggaagctctctgagggagctgatatgtatgcagaccagctcaag         | 839  |
| Db | 361  | ggtgtatctcacttggaagctctctgagggagctgatatgtatgcagaccagctcaag         | 420  |
| Oy | 840  | gggaaagaagatgtataaagctctccctcgctggaggagccatacatatgtctgacgtctg      | 899  |
| Db | 421  | gggaaagaagatgtataaagctctccctcgctggaggagccatacatatgtctgacgtctg      | 480  |
| Oy | 900  | aagaagaatcgtgtccaatctggtcctcgaccacactggtccttaactctcaactcttcac      | 959  |
| Db | 481  | aagaagaatcgtgtccaatctggtcctcgaccacactggtccttaactctcaactcttcac      | 540  |
| Oy | 960  | gtcgacccttgtaaaagacttgaattccagcctcatttgagccctcactagtatgtagaaga     | 1019 |
| Db | 541  | gtcgacccttgtaaaagacttgaattccagcctcatttgagccctcactagtatgtagaaga     | 600  |
| Oy | 1020 | gggagctcgtgccaaagaaagacacagacactcttgacacaaattataactctttgcgtta      | 1079 |
| Db | 601  | gggagctcgtgccaaagaaagacacagacactcttgacacaaattataactctttgcgtta      | 660  |
| Oy | 1080 | tttgatgaaggagaaagcttggtgcctcagaaacaaagacactctctgtagcagatagagat     | 1139 |
| Db | 661  | tttgatgaaggagaaagcttggtgcctcagaaacaaagacactctccctcagcaagatagagat   | 720  |
| Oy | 1140 | gctgcatactgtctcgggcctggtgcctaaatgycacacagtcacatgtgtatgttaacagctc   | 1199 |
| Db | 721  | gctgcatactgtctcgggcctggtgcctaaatgycacacagtcacatgtgtatgttaacagcgc   | 780  |
| Oy | 1200 | ctgcacagctgtgttttgatgtcccaaggaataatcattatgttgacatgttgatggatgggc    | 1259 |
| Db | 781  | ctgcacagctgtgttttgatgtcccaaggaataatcattatgttgacatgttgatggatgggc    | 840  |
| Oy | 1260 | accactctcgaagtgacatacaatatctctcgaaggtcacacatatctctgtgagaacacat     | 1319 |
| Db | 841  | accactctcgaagtgacatacaatatctctcgaaggtcacacatatctctgtgagaacacat     | 900  |
| Oy | 1320 | cgccagggcgtctcttggaatactctgcacaaataacttctccttaactgtcacaacactttgatg | 1379 |
| Db | 901  | cgccagggcgtctcttggaatactctgcacaaataacttctccttaactgtcacaacactctcatg | 960  |
| Oy | 1380 | gaccttggtgacagtttctactggtttgtcatatctcttcccaacaaatgtagatggaa        | 1439 |
| Db | 961  | gaccttggtgacagtttctactggtttgtcatatctcttcccaacaaatgtagatggaa        | 1020 |
| Oy | 1440 | gcttatgtcaaaagttagacagctgtgccagaggaaaccccaactacgatacgtaaaaataatgaa | 1499 |
| Db | 1021 | gcttatgtcaaaagttagacagctgtgccagaggaaaccccaactacgatacgtaaaaataatgaa | 1080 |
| Oy | 1500 | gaagcggaaagactatgatgtgatacttctactcgtattctgaatgtgattggtcaggtttgat   | 1559 |
| Db | 1081 | gaagcggaaagactatgatgtgatacttctactcgtattctgaatgtgattggtcaggtttgat   | 1140 |
| Oy | 1560 | gattgaacactctccctccttataccaaattcgtccagcttgccaaagaagatccctaaact     | 1619 |
| Db | 1141 | gattgaacactctccctccttataccaaattcgtccagcttgccaaagaagatccctaaact     | 1200 |
| Oy | 1620 | tggtgatacatatactgtcgtcgtgaaaggagagacttgtagacatgctcccttagtccctgc    | 1679 |
| Db | 1201 | tggtgatacatatactgtcgtcgtgaaaggagagacttgtagacatgctcccttagtccctgc    | 1260 |
| Oy | 1680 | cccgatgacagaagattataaagtccaattttgaacaaatggccctcagcggattggttag      | 1739 |
| Db | 1261 | cccgatgacagaagattataaagtccaattttgaacaaatggccctcagcggattggttag      | 1320 |

|    |      |   |      |
|----|------|---|------|
| QY | 1740 | aagfacacaaaagccgactttatggatgaacagatgaacactttaaagcttgaagct             | 1799 |
| Db | 1321 | aagfacacaaaagccgactttatggatgaacacagatgaacactttaaagcttgaagct           | 1380 |
| QY | 1800 | attcaagatgaatcaagaaatcttggagaccttactttaaggaggatcttggagacacatg         | 1859 |
| Db | 1381 | attcagcatgatcatgaagaaactcttggagaccttactttaaggaggatcttggagacacatg      | 1440 |
| QY | 1860 | ttgattatatttaagaatcaagcaagcgagacatataaactactaccctcaggaatcact          | 1919 |
| Db | 1441 | ctctatattatlaagaatcaagcaagcgagacatataaactactaccctcaggaatcact          | 1500 |
| QY | 1920 | gatgtccgtctcttggatctcaaggaattacaaccaaggtgttaaacatttgaagatttt          | 1979 |
| Db | 1501 | gatgtccgtctcttggatctcaaggaattacaaccaaggtgttaaacatttgaagatttt          | 1560 |
| QY | 1980 | ccaattctgcagagagaataatatacaataatgacagatgactgtgtaagaagtgtggcca         | 2039 |
| Db | 1561 | ccaattctgcagagagaataatatacaataatgacagatgactgtgtaagaagtgtggcca         | 1620 |
| QY | 2040 | actataatcaaatctcgtgtgcctgaccgcgtatctactagtttgcgttaaatgtagaga          | 2099 |
| Db | 1621 | actataatcaaatctcgtgtgcctgaccgcgtatctactagtttgcgttaaatgtagaga          | 1680 |
| QY | 2100 | gatactagcttcagagatcattctgtgcctctctctcatctgtctaaagaagaaactgtatgataa    | 2159 |
| Db | 1681 | gatactagcttcagagatcattctgtgtgcctctctctcatctgtctaaagaagaaactgtatgataa  | 1740 |
| QY | 2160 | agagagaacccaagataatgctcagaacagaggaatgtcatctcgttttctcgtatttggatgag     | 2219 |
| Db | 1741 | agagagaacccaagataatgctcagaacagaggaatgtcatctcgttttctcgtatttggatgag     | 1800 |
| QY | 2220 | aaccggaagctgtgtacccctccacaagaaatatacaacgcttctctcccaatccagctgtgaatg    | 2279 |
| Db | 1801 | aaccggaagctgtgtacccctccacaagaaatatacaacgcttctctcccaatccagctgtgaatg    | 1860 |
| QY | 2280 | cagctgtgagatgccagaagtctcaagcctccacaatcatgycacagacatcaatgtgctatgct     | 2339 |
| Db | 1861 | cagctgtgagatgccagaagtctcaagcctccacaatcatgycacagacatcaatgtgctatgct     | 1920 |
| QY | 2340 | tttatagtttgcagatctgttcagtttttggacagaggtgtgcatactgtgttaacttaagc        | 2399 |
| Db | 1921 | tttatagtttgcagatctgttcagtttttggacagaggtgtgcatactgtgttaacttaagc        | 1980 |
| QY | 2400 | attgagacagacagactgactctctcttctgtctctctctcgtgtatatacctcacaacaaa        | 2459 |
| Db | 1981 | attgagacagacagactgactctctcttctgtctctctctcgtgtatatacctcacaacaaa        | 2040 |
| QY | 2460 | atggtctatgaagacaacactcacccctatcccatctctcagggaaacgttctatgctg           | 2519 |
| Db | 2041 | atggtctatgaagacaacactcacccctatcccatctctcagggaaacgttctctatgctg         | 2100 |
| QY | 2520 | atggaataaacccaagctctatgtgattctctgggtgtgcacaacatcagaaattctcggaaacagagc | 2579 |
| Db | 2101 | atggaataaacccaagctctatgtgattctctgggtgtgcacaacatcagaaattctcggaaacagagc | 2160 |
| QY | 2580 | atgacgccttactgtaaggttctctagttgtgtgacagaagacacactgtgtatatacgaagac      | 2639 |
| Db | 2161 | atgacgccttactgtaaggttctctagttgtgtgacagaagacacactgtgtatatacgaagac      | 2220 |
| QY | 2640 | agttatgaagaatatcttcagagatactctgtcgtgagttaaanaaatgtcatcttggaaacgaagac  | 2699 |
| Db | 2221 | agttatgaagaatatcttcagagatactctgtcgtgagttaaanaaatgtcatcttggaaacgaagac  | 2280 |
| QY | 2700 | ttctcccaagatccaagacaccccttagcacttaggcaaaagccaatttcaatgcacccca         | 2759 |
| Db | 2281 | ttctcccaagatccaagacaccccttagcacttaggcaaaagccaatttcaatgcacccca         | 2298 |
| QY | 2760 | gtctcttgaagaagccatacgaaggaataactcgttactactcttcagtcagatccaagagaa       | 2819 |
| Db | 2299 | gtctcttgaagaagccatacgaaggaataactcgttactactcttcagtcagatccaagagaa       | 2358 |
| QY | 2820 | atttaccatataatcacatatgaattgttgaatgaagaagaagaagaagattttgacattatgat     | 2879 |

|||||  
Db 2359 attgacatgatatgataccataatcagttgaatgaagaagatttcgacattatgat 2418  
QY 2880 gagatgaataatcagagcccccagcgttcaaaaaaacaagcactattttatgct 2939  
Db 2419 gagatgaataatcagagcccccagcgttcaaaaaaacaagcactattttatgct 2478  
QY 2940 gcaatggagagcctctggtatattagtgatgtagctcccaatggtcttaagaacagg 2999  
Db 2479 gcaatggagagcctctggtatattagtgatgtagctcccaatggtcttaagaacagg 2538  
QY 3000 gctcagagctggcagctgtccctcagcttcaagaagaatgtgttttcagaggaatttactgctg 3059  
Db 2539 gctcagagctggcagctgtccctcagcttcaagaagaatgtgttttcagaggaatttactgctg 2598  
QY 3060 tcccttactcagcccttaccgctggagaactaaatgaacatttgggaactctgggcca 3119  
Db 2599 tcccttactcagcccttaccgctggagaactaaatgaacatttgggaactctgggcca 2658  
QY 3120 tataaagaagcagaagttgaagataataatcatgtgaacttcagaagaatcaggccctcgt 3179  
Db 2659 tatataagaagcagaagttgaagataataatcatgtgaacttcagaagaatcaggccctcgt 2718  
QY 3180 cccatttcctctatattcagccttatttctatagagaagaatcagaagcagaagcagaa 3239  
Db 2719 cccatttcctctatattcctccatcatatcatatgagaagaatcagaagcagaagcagaa 2778  
QY 3240 ccttagaaaaaacttctgcagagcctaaatgaacccaacttcttggaaagtgcacacat 3299  
Db 2779 ccttagaaaaaacttctgcagagcctaaatgaacccaacttcttggaaagtgcacacat 2838  
QY 3300 catatggacaccactaaagaatgaagtttgagctgcaagcctggcctatttccgagt 3359  
Db 2839 catatggacaccactaaagaatgaagtttgagctgcaagcctggcctatttccgagt 2898  
QY 3360 gacctggaaaaaagtctgctcagctcagcctgaatggacccctctgcttcgacaactaac 3419  
Db 2899 gacctggaaaaaagtctgctcagcctcagcctgaatggacccctctgcttcgacaactaac 2958  
QY 3420 acactggaacccctgtctcatatggagaagaatgacagctacaggaatttgcctgttttcacc 3479  
Db 2959 acactggaacccctgtctcatatggagaagaatgacagctacaggaatttgcctgttttcacc 3018  
QY 3480 atctttgtagacccaagaatctgttacttcaactgaataatcgaagaagaactgcaggct 3539  
Db 3019 atctttgtagacccaagaatctgttacttcaactgaataatcgaagaagaactgcaggct 3078  
QY 3540 cccctgcaaatcagatggaagatcccaactttaagaagaattatcgcttcacatgcaatc 3599  
Db 3079 cccctgcaaatcagatggaagatcccaactttaagaagaattatcgcttcacatgcaatc 3138  
QY 3600 aatggtcaataatggaatgatacactacgtgcttaagtgaatgctcaggaatcgaagaatcga 3659  
Db 3139 aatggtcaataatggaatgatacactacgtgcttaagtgaatgctcaggaatcgaagaatcga 3198  
QY 3660 tggatctgctcagcagctgggagcaatgaataacatccatctatctatcttaagtgagcat 3719  
Db 3199 tggatctgctcagcagctgggagcaatgaataacatccatctatctatcttaagtgagcat 3258  
QY 3720 gtgttcaactgtacgaaaaaagaagagatataaaatggcactgtacaatctatccaggt 3779  
Db 3259 gtgttcaactgtacgaaaaaagaagagatataaaatggcactgtacaatctatccaggt 3318  
QY 3780 gtttttgaagcagtggaagaatgtttacatcccaagctggaatttggcggtggaatgctt 3839  
Db 3319 gtttttgaagcagtggaagaatgtttacatcccaagctggaatttggcggtggaatgctt 3378  
QY 3840 attgcgagacatcacaatgctcgagatgagcacacttttctggtgacagaataaagtgt 3899  
Db 3379 attgcgagacatcacaatgctcgagatgagcacacttttctggtgactcacaataaagtgt 3438  
QY 3900 cagaatccctgggaatggtctctggaacatatagagatttccagattcagccttcagga 3959  
|||||

Db 3439 cagactccctgggaatggtctctggaacatatagagaatttcagattacagcttcagga 3498  
QY 3960 caataatgaacagtggtgcccccaagaatggtgcagacttaattatccggatcaatcatgcc 4019  
Db 3499 caataatgaacagtggtgcccccaagaatggtgcagacttaattatccggatcaatcatgcc 3558  
QY 4020 tggagcacaagaagcccttctctggaatgaagtggaatcgttttggaacccaatgattat 4079  
Db 3559 tggagcacaagaagcccttctctggaatgaagtggaatcgttttggaacccaatgattat 3618  
QY 4080 caggcatcacaagaacccaggctgcctcagagaagttctccaagcctcatalactcctcagtt 4139  
Db 3619 caggcatcacaagaacccagggtgcccgtcagaagatctccagcctcatalactcctcagtt 3678  
QY 4140 atcatcatgtatagtcttgaatgggaagaagtggcagacttatcgagaagaattccactgga 4199  
Db 3679 atcatcatgtatagtcttgaatgggaagaagtggcagacttatcgagaagaattccactgga 3738  
QY 4200 accttaatggtctctcttggcaatgtgattcatcttggaataaaacacaatatatttttac 4259  
Db 3739 acctcctagtgctctcttggcaatgtgattcatcttggaataaaacacaatatatttttac 3798  
QY 4260 cctccaattatgtctcgatataatccgtttgcaaccaactatataagatttcgacagcat 4319  
Db 3799 cctccaattatgtctcgatataatccgtttgcaaccaactatataagatttcgacagcat 3858  
QY 4320 ctctgcagtgagtgatggtgctgtgatttaaatagttgcagatgcatttgggaatggag 4379  
Db 3859 ctctgcagtgagtgatggtgctgtgatttaaatagttgcagatgcatttgggaatggag 3918  
QY 4380 agtaagaacatatcagatgcagacaagattactgtccatccctacttaccataatgtttgcc 4439  
Db 3919 agtaagaacatatcagatgcagacaagattactgtccatccctacttaccataatgtttgcc 3978  
QY 4440 acctgtctccttcaaaaagctcagactcaactccaaggaggaatgaatgcttggagacct 4499  
Db 3979 acctgtctccttcaaaaagctcagactcaactccaaggaggaatgaatgcttggagacct 4038  
QY 4500 caggtgaataatccaaaagaatgtgctgaagtggattccagaagacaatgaagaatcaca 4559  
Db 4039 caggtgaataatccaaaagaatgtgctgaagtggattccagaagacaatgaagaatcaca 4098  
QY 4560 ggaatgaactcagagtggaataaactctgtcttaacagcatgtatgtgaagagtctctc 4619  
Db 4099 ggaatgaactcagagtggaataaactctgtcttaacagcatgtatgtgaagagtctctc 4158  
QY 4620 atctccagcagtcagaatggtccatcagtggaactcctttttcagaatggtcaagaatgaag 4679  
Db 4159 atctccagcagtcagaatggtccatcagtggaactcctttttcagaatggtcaagaatgaag 4218  
QY 4680 gtttttcaggaataatcaagaactccttcaacactgtgtggaactctctagaccacagctta 4739  
Db 4219 gtttttcaggaataatcaagaactccttcaacactgtgtggaactctctagaccacagctta 4278  
QY 4740 ctgactcgtactacactcagatccaccccagaagttgggtgacccagatggccttgaagatt 4799  
Db 4279 ctgactcgtactacactcagatccaccccagaagttgggtgacccagatggccttgaagatt 4338  
QY 4800 gaggttctggctcgagagcacaaggaactctactg 4834  
Db 4339 gaggttctggctcgagagcacaaggaactctactg 4373  
  
RESULT 15  
AANB1544  
ID AANB1544 standard; DNA; 4830 BP.  
XX  
XX  
AC AANB1544;  
XX  
DT 04-DEC-1990 (first entry)  
XX  
DE Human Factor VIII-C analog having exon 14 deleted.  
XX

|    |   |
|----|---|
| KW | Human Factor VIII-C analog; exon deletion; coagulation disorders.   |
| KX | haemophilia; ss.  |
| XX |   |
| PN | EP265778-A.   |
| XX |   |
| PD | 04-MAY-1988.  |
| XX |   |
| PF | 14-OCT-1987; 87EP-0115043.  |
| XX |   |
| PR | 15-OCT-1986; 86US-0919153.  |
| XX |   |
| PA | (RORE ) KOKER INT OVERSEAS.   |
| XX |   |
| PI | Sarver N, Drohan W;   |
| XX |   |
| DR | WPI. 1988-120930/18.  |
| XX |   |
| PT | Human Factor VIII-C analogue free of other proteins - produced from |
| PT | recombinant DNA and used in the treatment of coagulation disorders. |
| XX |   |
| PS | Clatm 3; Page 13; 42pp; English.                                    |
| XX |   |
| CC | The genetically engineered analogue can provide a dependable and    |
| CC | readily available therapeutic agent to be used in the treatment of  |
| CC | haemophilia and coagulation disorders in humans or animals.         |
| CC | See also AAN81543-45.   |
| XX |   |
| SO | Sequence 4830 BP; 1392 A; 1091 C; 1058 G; 1289 T; 0 other;          |

|    |      |  |      |
|----|------|--|------|
| QY | 2100 | gactagcttcaggaactcaatctggccctccctccatctcgtctcaagaagatctgtatgacaa   | 2159 |
| Db | 1681 | gacttagctctcaggaactcaatctggccctccctccatctcgtctcaagaagatctgtatgacaa | 1740 |
| QY | 2160 | agaggaacacataatctgcacagaggaagtgcatacctgtcttctgtatttgatgag          | 2219 |
| Db | 1741 | agaggaacacataatctgcacagaggaagtgcatacctgtcttctgtatttgatgag          | 1800 |
| QY | 2220 | aaccgaagcttgcacctccacagaaatatacaacgcttctcccaatccagcttgagtg         | 2279 |
| Db | 1801 | aaccgaagcttgcacctccacagaaatatacaacgcttctcccaatccagcttgagtg         | 1860 |
| QY | 2280 | caagcttgagatccagatgtccaaagcctccacaatcatatgcagaacatctggtatgtt       | 2339 |
| Db | 1861 | caagcttgagatccagatgtccaaagcctccacaatcatatgcagaacatctggtatgtt       | 1920 |
| QY | 2340 | tttgatagtttcagatgttcagattgtttgcagaggttgcatccggtgacattctcaagc       | 2399 |
| Db | 1921 | tttgatagtttcagatgttcagattgtttgcagaggttgcatccggtgacattctcaagc       | 1980 |
| QY | 2400 | attggagacacagactgactctctctctgtctctctccatgcagagaaactgctctcatgctg    | 2459 |
| Db | 1981 | attggagacacagactgactctctctctgtctctctccatgcagagaaactgctctcatgctg    | 2040 |
| QY | 2460 | atggtctatgaaagacacactccactccattcccatctccagagaaactgctctcatgctg      | 2519 |
| Db | 2041 | atggtctatgaaagacacactccactccattcccatctccagagaaactgctctcatgctg      | 2100 |
| QY | 2520 | atgggaacacccaggtctatggaattctcggtgtgcacaactccagacttctcggaacagagc    | 2579 |
| Db | 2101 | atgggaacacccaggtctatggaattctcggtgtgcacaactccagacttctcggaacagagc    | 2160 |
| QY | 2580 | atgacccgcttaactggaagttctcaattgttgacaagaacacgtgtatctatctacagagag    | 2639 |
| Db | 2161 | atgacccgcttaactggaagttctcaattgttgacaagaacacgtgtatctatctacagagag    | 2220 |
| QY | 2640 | agttatgaaagatatctcagacatactgtcgtgaaataaacaatgcatctgaaaccaagaagc    | 2699 |
| Db | 2221 | agttatgaaagatatctcagacatactgtcgtgaaataaacaatgcatctgaaaccaagaagc    | 2280 |
| QY | 2700 | ttctccgaatctca-----ttctccgaatctca-----                             | 2714 |
| Db | 2281 | ttctccgaatctcaatctatctgatacctctgtcttgagataaccaatctgatacagata       | 2340 |
| QY | 2715 | -----  | 2714 |
| Db | 2341 | ccaaaaagaagatggaatcccaagaagaatcccaagaagaacagctttaaagaagaag         | 2400 |
| QY | 2715 | -----  | 2714 |
| Db | 2401 | gatacaatttgccttgacgtgtgtgaagaacatcatgcaatagcagcaataatagtg          | 2460 |
| QY | 2715 | -----  | 2714 |
| Db | 2461 | ggacaaataagcccgaaatagaagtcaacctggtgcacagcaaggtgacatgaaagcgctg      | 2520 |
| QY | 2742 | tttaattgcaccccccagctctctgaaagccatcaacgaggaataactcgtatactctt        | 2801 |
| Db | 2521 | ttgcttcaaaccccaacgctctgaaagccatcaacgaggaataactcgtatactctt          | 2580 |
| QY | 2802 | caatcagatcaagaagaaatctgactatgatacatalcagcttgtaaaatgaagaagaa        | 2861 |
| Db | 2581 | caatcagatcaagaagaaatctgactatgatacatalcagcttgtaaaatgaagaagaa        | 2640 |
| QY | 2862 | gattttgataattatgataagatatgaataatcagagcccccagcgttcaaaagaanaa        | 2921 |
| Db | 2641 | gattttgataattatgataagatatgaataatcagagcccccagcgttcaaaagaanaa        | 2700 |
| QY | 2922 | caaacactatttatgctcagctgagagagccttgatattatgagtgagtgagctccca         | 2981 |
| Db | 2701 | caaacactatttatgctcagctgagagagccttgatattatgagtgagtgagtgagctccca     | 2760 |
| QY | 2982 | catgttctaagaacagggctcagagtgagctgtccctcagtttcaagaagaattgttttc       | 3041 |





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 17, 2001, 17:14:18 ; Search time 22045.3 Seconds  
(without alignments)  
5573.794 Million cell updates/sec

Title: US-09-689-430-1  
Perfect score: 7944  
Sequence: 1 tggccactccctctctgcgc.....agggggggggggggggggt 7944

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba1:\*  
2: gb\_ba2:\*  
3: gb\_ba3:\*  
4: gb\_in1:\*  
5: gb\_in2:\*  
6: gb\_in3:\*  
7: gb\_om:\*  
8: gb\_ov:\*  
9: gb\_pat1:\*  
10: gb\_pat2:\*  
11: gb\_ph:\*  
12: gb\_pl1:\*  
13: gb\_pl2:\*  
14: gb\_pl3:\*  
15: gb\_pl4:\*  
16: em\_ba1:\*  
17: em\_ba2:\*  
18: em\_fun:\*  
19: em\_htgo\_hum:\*  
20: em\_htgo\_iny:\*  
21: em\_htgo\_rod:\*  
22: em\_htg\_hum1:\*  
23: em\_htg\_hum2:\*  
24: em\_htg\_hum3:\*  
25: em\_htg\_hum4:\*  
26: em\_htg\_hum5:\*  
27: em\_htg\_hum6:\*  
28: em\_htg\_hum7:\*  
29: em\_htg\_hum8:\*  
30: em\_htg\_invl:\*  
31: em\_htg\_invl2:\*  
32: em\_htg\_rod:\*  
33: em\_htg\_rod2:\*  
34: em\_hum1:\*  
35: em\_hum2:\*  
36: em\_hum3:\*  
37: em\_hum4:\*  
38: em\_hum5:\*  
39: em\_hum6:\*  
40: em\_hum7:\*  
41: em\_in:\*  
42: em\_om:\*  
43: em\_or:\*

44: em\_ov:\*  
45: em\_pat:\*  
46: em\_ph:\*  
47: em\_pl:\*  
48: em\_ro:\*  
49: em\_sts:\*  
50: em\_sy:\*  
51: em\_un:\*  
52: em\_vi:\*  
53: gb\_sts1:\*  
54: gb\_sts2:\*  
55: gb\_sts3:\*  
56: gb\_sy:\*  
57: gb\_un:\*  
58: gb\_v11:\*  
59: gb\_v12:\*  
60: gb\_htg1:\*  
61: gb\_htg2:\*  
62: gb\_htg3:\*  
63: gb\_htg4:\*  
64: gb\_htg5:\*  
65: gb\_htg6:\*  
66: gb\_htg7:\*  
67: gb\_htg8:\*  
68: gb\_htg9:\*  
69: gb\_htg10:\*  
70: gb\_htg11:\*  
71: gb\_htg12:\*  
72: gb\_htg13:\*  
73: gb\_htg14:\*  
74: gb\_htg15:\*  
75: gb\_htg16:\*  
76: gb\_htg17:\*  
77: gb\_htg18:\*  
78: gb\_htg19:\*  
79: gb\_htg20:\*  
80: gb\_htg21:\*  
81: gb\_htg22:\*  
82: gb\_htg23:\*  
83: gb\_htg24:\*  
84: gb\_htg25:\*  
85: gb\_pr1:\*  
86: gb\_pr2:\*  
87: gb\_pr3:\*  
88: gb\_pr4:\*  
89: gb\_pr5:\*  
90: gb\_pr6:\*  
91: gb\_pr7:\*  
92: gb\_pr8:\*  
93: gb\_pr9:\*  
94: gb\_fo1:\*  
95: gb\_fo2:\*  
96: gb\_in4:\*  
97: gb\_pr10:\*  
98: em\_ba3:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID      | Description       |
|------------|--------|-------------|--------|------------|-------------------|
| 1          | 4411.6 | 55.5        | 9354   | 9 AR003585 | AR003585 Sequence |
| 2          | 4357.6 | 54.9        | 4670   | 9 AR110040 | AR110040 Sequence |
| 3          | 4148.8 | 52.2        | 4278   | 10 I08644  | I08644 Sequence 4 |
| 4          | 4147.2 | 52.2        | 4281   | 10 I08643  | I08643 Sequence 3 |
| 5          | 4085   | 51.4        | 4548   | 10 I08642  | I08642 Sequence 2 |
| 6          | 4085   | 51.4        | 4551   | 10 I08641  | I08641 Sequence 1 |
| 7          | 3748   | 47.2        | 5035   | 9 AR034084 | AR034084 Sequence |
| 8          | 3748   | 47.2        | 5035   | 9 AR071306 | AR071306 Sequence |

|    |        |      |       |    |           |                     |
|----|--------|------|-------|----|-----------|---------------------|
| 9  | 3189.8 | 40.2 | 4334  | 5  | AR029098  | AR029098 Sequence   |
| 10 | 2830.6 | 35.6 | 6786  | 56 | CVU89927  | AR99927 Cloning vec |
| 11 | 2809.4 | 35.4 | 3179  | 56 | CGEM132P  | X65315 Cloning vec  |
| 12 | 2809.4 | 35.4 | 3197  | 56 | CVPGEM32P | X65306 Cloning vec  |
| 13 | 2809.4 | 35.4 | 3221  | 56 | CGEM112P  | X65313 Cloning vec  |
| 14 | 2794.2 | 35.2 | 4118  | 9  | AS9363    | AS9363 Sequence 17  |
| 15 | 2794.2 | 35.2 | 4118  | 9  | AS9364    | AS9364 Sequence 18  |
| 16 | 2781   | 35.0 | 3582  | 56 | SYNEOMCS  | M68846 Cloning vec  |
| 17 | 2781   | 35.0 | 3642  | 56 | SYN260BLA | M29362 Plasmid pET  |
| 18 | 2764.2 | 34.8 | 4076  | 56 | AF013258  | M29362 Expressio    |
| 19 | 2764.2 | 34.8 | 4905  | 56 | ASPGST    | X78316 Artificial   |
| 20 | 2762.6 | 34.8 | 6094  | 9  | AS6973    | AS6973 Sequence 3   |
| 21 | 2758.8 | 34.7 | 3984  | 10 | 115353    | 115353 Sequence 1   |
| 22 | 2741.8 | 34.5 | 5676  | 56 | CVSELECT  | X63334 Cloning vec  |
| 23 | 2719   | 34.2 | 5858  | 56 | CVU47102  | U47102 Cloning vec  |
| 24 | 2699.2 | 34.0 | 6101  | 56 | CVPSIT    | U17139 Phagemid cl  |
| 25 | 2637.8 | 33.2 | 3162  | 10 | E14304    | U14304 pUC118. 7/1  |
| 26 | 2637.8 | 33.2 | 8532  | 56 | U02452    | U02452 Cloning vec  |
| 27 | 2637.8 | 33.2 | 9595  | 10 | E16678    | E16678 All sequenc  |
| 28 | 2636.2 | 33.2 | 4696  | 9  | AR069374  | AR069374 Sequence   |
| 29 | 2636.2 | 33.2 | 5158  | 9  | AR069375  | AR069375 Sequence   |
| 30 | 2636.2 | 33.2 | 7607  | 10 | 144516    | 144516 Sequence 19  |
| 31 | 2636.2 | 33.2 | 19307 | 9  | AR093392  | AR093392 Sequence   |
| 32 | 2631.6 | 33.1 | 5461  | 56 | U02450    | U02450 Cloning vec  |
| 33 | 2631.6 | 33.1 | 5538  | 56 | U02446    | U02446 Cloning vec  |
| 34 | 2631.6 | 33.1 | 5538  | 56 | XXU02440  | U02440 Cloning vec  |
| 35 | 2631.6 | 33.1 | 11958 | 9  | AX057269  | AX057269 Sequence   |
| 36 | 2631.4 | 33.1 | 9208  | 56 | AF064067  | AF064067 Expressio  |
| 37 | 2631.4 | 33.1 | 9211  | 56 | AF064066  | AF064066 Expressio  |
| 38 | 2631.4 | 33.1 | 9215  | 56 | AF049063  | AF049063 Expressio  |
| 39 | 2631.4 | 33.1 | 9956  | 56 | AF049064  | AF049064 Expressio  |
| 40 | 2625.8 | 33.1 | 4054  | 9  | A81878    | A81878 Sequence 9   |
| 41 | 2621   | 33.0 | 3632  | 56 | CVU47121  | U47121 Cloning vec  |
| 42 | 2621   | 33.0 | 4006  | 56 | CVU47119  | U47119 Cloning vec  |
| 43 | 2605.8 | 32.8 | 4910  | 56 | ASAJ156   | AJ000156 Artificial |
| 44 | 2605.8 | 32.8 | 4951  | 56 | ASAJ154   | AJ000154 Artificial |
| 45 | 2605.8 | 32.8 | 5284  | 56 | ASOURESN  | Y11035 Artificial   |

## ALIGNMENTS

|            |   |            |     |     |             |  |
|------------|---|------------|-----|-----|-------------|--|
| RESULT     | 1   |            |     |     |             |  |
| LOCUS      | AR003585  |            |     |     |             |  |
| DEFINITION | AR003585  | 9354 bp    | DNA | PAT | 04-DEC-1998 |  |
| ACCESSION  | Sequence 2 from patent US 5744326.  |            |     |     |             |  |
| VERSION    | AR003585  |            |     |     |             |  |
| KEYWORDS   | AR003585.1  | GI:3964844 |     |     |             |  |
| SOURCE     | .   |            |     |     |             |  |
| ORGANISM   | Unknown.  |            |     |     |             |  |
|            | Unknown.  |            |     |     |             |  |
|            | Unclassified.   |            |     |     |             |  |
| REFERENCE  | 1 (bases 1 to 9354)   |            |     |     |             |  |
| AUTHORS    | Ill,C.R. and Blottingmaler,S.   |            |     |     |             |  |
| TITLE      | Use of viral CIs-acting post-transcriptional regulatory sequences<br>to increase expression of intronless genes containing<br>near-consensus splice sites |            |     |     |             |  |
| JOURNAL    | Patent: US 5744326-A 2 28-Apr-1998;   |            |     |     |             |  |
| FEATURES   | Location/Qualifiers   |            |     |     |             |  |
| source     | 1..9354   |            |     |     |             |  |
|            | /organism="unknown"   |            |     |     |             |  |
| BASE COUNT | 2506 a 2239 c 2161 g 2448 t   |            |     |     |             |  |
| ORIGIN     |   |            |     |     |             |  |

|    |                       |  |               |                |              |         |
|----|-----------------------|--|---------------|----------------|--------------|---------|
|    | Query Match           | 55.5%  | Score 4411.6: | DB 9:          | Length 9334: |         |
|    | Best Local Similarity | 99.7%:   | Pred. No. 0:  |                |              |         |
|    | Matches 4420:         | Conservative   | 0:            | Mismatches 14: | Indels 0:    | Gaps 0: |
| OY | 415                   | ccacatgcgaataagactcctcacotggttccttctcgaccttttgacatcctgltta | 474           |                |              |         |
|    |                       |  |               |                |              |         |
| Db | 2960                  | CCACATGATTATAGACTCTCACACTCGTTTCTTCGTGGCCTTTTGACATTCGCTTFA  | 3019          |                |              |         |

|    |      |  |      |
|----|------|--|------|
| QY | 475  | gfhccacccaagaagatctactcttggtgtccagtgtgaacgctcaatgtgagctcttatgcaaaag    | 534  |
| Db | 3020 | gtggccaccagaagatctactcttggtgtccagtgtgaacgctcaatgtgagctcttatgcaaaag     | 3078 |
| QY | 535  | alctcgtgtgaagctgcctgtgtgaagccaagatctccctcttagatgccaatactcttccat        | 594  |
| Db | 3080 | attctcgtgtgaagctgcctgtgtgaagccaagatctccctcttagatgccaatactcttccat       | 3138 |
| QY | 555  | lcaaacactcagtcgtgttacaaaaagacctcgtttttagaatcaacagtttcaactttca          | 654  |
| Db | 3140 | tcaaacactcagtcgtgtgtgtacaaaaagacctcgtttttagaatcaacagtttcaactttca       | 3198 |
| QY | 655  | acatgcctcaagccaagccaccctcgatgtgtctcgaagtcctaccatcccgagctcgag           | 714  |
| Db | 3200 | acatgcctcaagccaagccaccctcgatgtgtctcgaagtcctaccatcccgagctcgag           | 3258 |
| QY | 715  | tttatgatacagtggttcaatcacttaagacaatggttcccactcctgtccagtctcatg           | 774  |
| Db | 3260 | ttttatgatacagtggttcaatcacttaagacaatggttcccactcctgtccagtctcatg          | 3318 |
| QY | 775  | ctgtgtgtgtatctcctactgtgaagactcttgaaggagctgaatatgtatgcagaccagtc         | 834  |
| Db | 3320 | ctgtgtgtgtatctcctactgtgtgtgaagactcttgaaggagctgaatatgtatgcagaccagtc     | 3378 |
| QY | 835  | aaagggagaaaagaatgataaagctctctccctgtgtgtgaagcaatactgtctggcag            | 894  |
| Db | 3380 | aaagggagaaaagaatgataaagctctctccctgtgtgtgtgaagcaatactgtctggcag          | 3438 |
| QY | 895  | tcctbaaagagaatggtgtccaaatgtggcctctgaccacacgtgtccttacctactaatctt        | 954  |
| Db | 3440 | tcctbaaagagaatggtgtgtgtccaaatgtggcctctgaccacacgtgtccttacctactaatctt    | 3498 |
| QY | 955  | ctcatgtgagactctgtgtlaaaaagactctgaatctcaagcctcatcttgagagcctactagtatla   | 1014 |
| Db | 3500 | ctcatgtgagactctgtgtgtlaaaaagactctgaatctcaagcctcatcttgagagcctactagtatla | 3558 |
| QY | 1015 | gagaaagagatctctggccaaggaagaagacaagactctgccaatacttaactactctttg          | 1074 |
| Db | 3560 | gagaaagagatctctggccaaggaagaagacaagactctgccaatacttaactactctttg          | 3618 |
| QY | 1075 | ctgtatcttgatggaagggaagaaagtgtggaacccagaacaagaactcctctgtatgacagata      | 1134 |
| Db | 3620 | ctgtatcttgatggaagggaagaaagtgtggaacccagaacaagaactcctctgtatgacagata      | 3678 |
| QY | 1135 | ggagatctgtcactctgtcctcgtggccttgycctlaaaatgtcacacagtcacaaatgtgtatgaaca  | 1194 |
| Db | 3680 | ggagatctgtcatctgtcctcgtggccttgycctlaaaatgtcacacagtcacaaatgtgtatgaaca   | 3738 |
| QY | 1195 | ggtctctgcacagctctgtatgtgatgtgccaagaagatcagtcactgtgtgcattgtgtga         | 1254 |
| Db | 3740 | ggtctctgcacagctctgtatgtgatgtgccaagaagatcagtcactgtgtgcattgtgtga         | 3798 |
| QY | 1255 | tggggacacacccctgaaggtgacacaaatctcctcgaaggtcaacaaatctctgtgagga          | 1314 |
| Db | 3800 | tggggacacacccctgaaggtgacacaaatctcctcgaaggtcaacaaatctctgtgagga          | 3858 |
| QY | 1315 | accatcgcgcaggtcctctgtgaaatctcgcgcacaaacttctccttactgttcaacaactc         | 1374 |
| Db | 3860 | accatcgcgcaggtcctctgtgaaatctcgcgcacaaacttctccttactgttcaacaactc         | 3918 |
| QY | 1375 | tgatgtgacctgtgacagttctctactgttttgtcatactctctcccaacaatgtatgagca         | 1434 |
| Db | 3920 | tgatgtgacctgtgtgacagttctctactgttttgtcatactctctcccaacaatgtatgagca       | 3978 |
| QY | 1435 | ttggaagcttatagtcaaagtgtgacagctgtctccagaagaaacccaactcaagaaatgtaaaaata   | 1494 |
| Db | 3980 | ttggaagcttatagtcaaagtgtgacagctgtctccagaagaaacccaactcaagaaatgtaaaaata   | 4038 |
| QY | 1495 | atggaagaagcggagaactatgatatgatatcttactgtactgtgaatgtgagtgtgtgcaggt       | 1554 |
| Db | 4040 | atggaagaagcggagaactatgatatgatatcttactgtactgtgaatgtgagtgtgtgcaggt       | 4098 |
| QY | 1555 | ttgatgtatgagaacactctcctccttcttaccaaaatcgcgtcagttgtgcagaagaacactcta     | 1614 |

|||||  
Db 4100 TTGATGATGACAACCTCTCTCTTATTCACAATTCGCTCAAGTGCCTCAAGATCCGA 4159  
OY 1615 aaacttggtaacttaacttctgtctgaagaggaagcttggacatctctcccttaagtc 1674  
Db 4160 AAACCTGGGTACATTAATTCCTGCTGAAGAGAGACTGGGACATATGCTCCTTAGGCC 4219  
OY 1675 tgcgccccgatcacagaagttataaagttcaattttgaacatggccctcagagatg 1734  
Db 4220 TCGCCCCGATACAGAAAGTTTAAAGTCAATATTGACACATGCCCCCTCAGCGGATTG 4279  
OY 1735 gtaaggaattacaaaaaagttccgaattatgycacacagatgaaaccttlaagactcg 1794  
Db 4280 GTAGGAAGTACAAAAGATCCGATTTATGCGATACACAGATGAACCTTTAAGACTCGG 4339  
OY 1795 aagcattcagatgaatcacgaatctctgggaaccttacttcttaagggaagtgtgagca 1854  
Db 4340 AAGCTATTCAGATGAATCAGGAATCTTGGACCTTTACTTTATGGGAGAGTTGGAGACA 4399  
OY 1855 cactgtgattatatacttaagaatcaagaacagacacataacatctacccctcagaa 1914  
Db 4400 CACTGTGATTAATTAATTAAGAAATCAAGACAGACACATATACATCTACCTCAGGAA 4459  
OY 1915 tcaactgatgtccgtctctgtatccaagagatlaacaaaagtgtaaacatttgaag 1974  
Db 4460 TCACGTAGTCCGCTCTTTGATTCAGAGAGATACAAAAGTGTAACATTTGAAGG 4519  
OY 1975 atttcacattctgcagagagaataatccaaataataatgacagtgactgtagaagatg 2034  
Db 4520 ATTTTCCAAATTCGACAGAGAAATTAATTAATGAACATGACTGTAGAGATG 4579  
OY 2035 ggcacaactaatcagatctcgtgcctgcacccgctatctacttactgttctgaatgag 2094  
Db 4580 GGCCAACTAATACAGATCTCGTGCTGACCCGCTATTACTCTAGTTTCCGTAATAGG 4639  
OY 2095 agaagatctagctcaggaactatggccctcctcactctgctacaaagaatctgtag 2154  
Db 4640 AGAGAGATCTGACTTCAGAGACTATGGCCCTCTCCTCATCTGATCAAGAAATCTGAG 4699  
OY 2155 atcaagagggaacacagtaattgtcagacaaagagaatgcatcctgttctctgatttg 2214  
Db 4700 ATCAAGAGGAAACCCAGATTAATGTCAGCAAGAGGAATGTCATCTGTTTCTGTAATTTG 4759  
OY 2215 atgaaacacgaagctgtgacctcacagagaatatacaacgcttctcccaatccagctg 2274  
Db 4760 ATGAACACCGAAGCTGTGTAACCTCAAGAGATATACACGCTTTTCCCCAATCCAGCTG 4819  
OY 2275 gaagtgcagcttgaagatccagaagttccaagctcccaacatcatgacagcaaccaatg 2334  
Db 4820 GAGTGCACCTTGAGGATTCAGAGTTCCAAAGCCTCCAMCATCATGCACAGCATATGCT 4879  
OY 2335 atgttttgatagtttgtaoatgtgtcagttgtttgcatgagatgagtcacatgtaacatc 2394  
Db 4880 ATGTTTGTGATAGTTTGCAAGTTGTCAGTTTGTTCATGAGAGTGCAATGTAATTC 4939  
OY 2395 taagcatggagacacagctgaactccttctcgtctctcctcctcgtgatatccctcaac 2454  
Db 4940 TTAAGATTGGAGACAGACTGACTTCTTCTGTCTTCTTCTGTGATATACCTTCAAC 4999  
OY 2455 acaaaatggtctatgaagacacacactacccatctcccaatctcagagagaatctgttca 2514  
Db 5000 ACAAAATGCTATGAAACACACACTCACCTTATCCATTCCTCAGAGAAACTGCTTCA 5059  
OY 2515 tgtcgaatgaaacacagttatagatctcgggtgtgcacaaactcagacttctggaaaca 2574  
Db 5060 TGTGATGGAAGAACCCAGCTCTATGAGATTCTGGGGTGCACAACTCAACATTTTGGAAACA 5119  
OY 2575 gaagcatgacccgtctactgaagttctctagttgtgacaagaacacatgtaattacag 2634  
Db 5120 GAGGATGACCGCTTACTGAAAGTTTCTAGTTGTGACAAACACCTGCTATTTATTCG 5179  
OY 2635 aggaacatgtaagatacttcagacatctgtcgtgagtaaaaacaatgcaattgaacca 2694  
|||||

Db 5180 AGACATGTTATGAAGATATTTACAGATCTTGTGAGTAAATAACATGCATTTGAACCA 5239  
OY 2695 gaagcttctcccaagaattccaagaacctagcaactagggcaaaagcaatttaatgccacc 2754  
Db 5240 GAAGCTTCTCCAGAAATTCAGACACCCCTACACAGCAAAAGCAATTTAATGCGACCC 5299  
OY 2755 caccagcttgaagacatcaacgggaataaactcgtactactccttcgaatcaatgaag 2814  
Db 5300 CACCGATTGGAACGCCATACCGGAAATTAACCTGACTACTCTTTCAGTCAGATCAG 5359  
OY 2815 aggaattgactatgatalatccatalcagttgaatgaagaagaatttgacatt 2874  
Db 5360 AGGAATGAGATGATGATACATATCAGTTGAATGAAGAGAGATTTTGACATTT 5419  
OY 2875 atgatgaagatgaanaatcaagggcccgcaacttccaagaagaacacacatcttla 2934  
Db 5420 ATGATGAGGATGAANAATGAGACCCCGCACCTTTCAAAAACACAGACATATTTTA 5479  
OY 2935 ttgctgcagtggaagggctcggggtatagagatgagctcccaatgttctaagaa 2994  
Db 5480 TTGCTGCAGTGAGAGGCTGTGGATTTATGGATGAGTACCTCCACATGTTCTGAAGA 5539  
OY 2995 acaaggctcagaatgtagcgtgtccctcagttcaagaagttgtttccaagaaattactg 3054  
Db 5540 ACAGGGCTCAGAGTGGCGAGTGTCCCTCAGTTCAAGAAAGTTGTTCCAGGAATTTACTG 5599  
OY 3055 atggtcctcttaccagcccttatacgttggaagactaaatggaactttggactccgtg 3114  
Db 5600 ATGGCTCTTTACTCAGGCCCTTATACCGTGAAGCACTAAATGAACATTTGGACCTCGG 5659  
OY 3115 ggcacatataagacagaatltgaagataatcatgtgaacttccaagaactcagacct 3174  
Db 5660 GGCCATATATAGACAGAGAAAGTTGAAGATATATCATGTGTAACCTTTCAAGATCAGGCT 5719  
OY 3175 ctgcctccatctccttacttactagacctatcttctlaagaaagatcagaagcaagag 3234  
Db 5720 CTCGTCCTATTCTCTTATCTAGCTTATTTCTTATGAGAGATCAGAGCAAGAG 5779  
OY 3235 cagaacctagaanaaacttgttcaagcctlaatgaaaccaaacttacttgggaagtgc 3294  
Db 5780 CAGAACCTGAGAAAACCTTGTCAAGCCTAATGAACCAAACTACTTTGGAAAGGC 5839  
OY 3295 aacatcatatgaccccaactaaagaatgagttgagctgcagaagcctgggctattctctg 3354  
Db 5840 AACATCATATGCGACCCCACTAAAGATAGTTGACTGCAAAAGCCTGATTTCTCTG 5899  
OY 3355 atgttgccttggaaaagaatgtgcaactcagctgattgagccctctgtgtgcaca 3414  
Db 5900 ATGTTGACCTGGAAGAAAGATGTGACATCAGGCTGATGAGACCCCTCTGTGTGCCACA 5959  
OY 3415 cttaacacactgaacccgtctcatgtggagacaaagtaagaagtaacgaattgtctgtttt 3474  
Db 5960 CTAACACACTGAACCCCTGCTATGAGACAAAGTACAGTACAGGAATTTGCTGTGTTT 6019  
OY 3475 tcaccatctctgatagagacaaagctgtgtaactcaactgaaacatgtgaagaacatgca 3534  
Db 6020 TCACACTTCTTGAATGAGACCAAAAGCTGTGCTTCTCAGTGAAGAAATGGAAGAAACGCA 6079  
OY 3535 gggctcccttgaatataccagaatgaagatccactttaaagagaattatgcttcatg 3594  
Db 6080 GGGCTCCCTGCAATATTCAGATGCAAGATCCCACTTTTAAAGAAATATATGCTTCCATG 6139  
OY 3595 caatcaatggtctataatagatatcacactcgtcttagtaatgtgtccaagatacaagga 3654  
Db 6140 CAATCAATGGCTATATATGATACACTGCTGCTTATGTAATGGCTCAGGATCAAGGA 6199  
OY 3655 ttgcagtgtactgtcgtcagcaatgggcaatgaaacatccatctatctatctcagtg 3714  
Db 6200 TTGCATGATATCTCTCAGCATGCGCAATGGAATGAAGATTCATTTATTTCACTG 6259  
OY 3715 gacatgttctcactgttcagaaaaaagaagatlaaaatgagcactgtacaaatctctatc 3774  
Db 6260 GACATGTTGTTCACTGTACGAAAAAGAGAGATTAATAATGGCACTGTACAACTCTATC 6319  
|||||

3775 caggtgttttttgagacagtgtaaatgttaccalccaaagcttgaaatttggcgggttgaaat 3834  
|||||  
6320 CAGGCTTTTTCAGACAGTGAAGTGTTCACATCCAAAGCTGGAATTTGGCGGTGGAAAT 6379  
3835 gacctatctggcgagacatcatcgtctggga tgaagcacactttctcgtgtatagacaata 3894  
|||||  
6380 GCGTATTGGCGAGCATCTACATGCTGGGATGAGCACACTTTTTCGGTGTACAGCAANA 6439  
3895 agtgcacaacccccctgggaatggctctctggaacacatagaatttcagatcaagctt 3954  
|||||  
6440 AGTGCACAGCTCCCTGGGAATGGCTTCTGACACATTGAGATTTTCAGATTACAGCTT 6499  
3955 caggacaaataggaacagtggtgccccaaagcttggccagacttcatatctccgaatcaca 4014  
|||||  
6500 CAGGACAAATATGACAGTGGGCCCCAAAGCTGGCCAGACTTCATTATTCGGATCAATCA 6559  
4015 atgcttggaagcaacgaagagcccttctcttgatcaaaagtgtgactcgttggcacaatga 4074  
|||||  
6560 ATGCTTGAGGACCAAGAGCCCTTTTCTTGATCAAGGTGATCTGTTGGCACCAATGA 6619  
4075 ttattcaggaatcaagaagccaggttgcccgctcaagaagttccagccctcatctctc 4134  
|||||  
6620 TTATTTCAGGGCTACAGACCCAGGGTCCCGTACAGAGTTCTCCAGCCTCTACATCTCTC 6679  
4135 agttatcatcatgatatagctcttga tgggaaagaagtgtgcagacttalcgaagaatcca 4194  
|||||  
6680 AGTTTATCATCATGTATAGTCTTGATGGGAGAGAAGTGGCAGACTTATTCAGGAAATTC 6739  
4195 ctggaaccttaagtctctcttggcaatgtgtatcatcttgggaataaaacacataatt 4254  
|||||  
6740 CTGGAACCTTATATGCTCTTCTTGGCAATGTGATTCATCTGGGATTAACCAATATTT 6799  
4255 ttaacccctcaatcttctcgaatcatccgttgcacccaactcttatacatctcga 4314  
|||||  
6800 TTAAACCTCCCAATTATTTGCTCGATACATCCGTTTGACCCCACTCTTATAGCATTCGA 6859  
4315 gcaactctcgaatgagtgatgtggcgtgtatataatagttgcagcaatgcatltggaa 4374  
|||||  
6860 GCACCTTCGCGATGAGTTGATGGGCTGTGATTTAAATAGTTGACAGCATGCCATTGGAA 6919  
4375 tggagagtaagaacatcatcagaatgtcaagataactgtctcatcttaccataatgt 4434  
|||||  
6920 TGGAGAGTAAGCAATATCAGATGACACAGATTACCTTCTTACCTTACCAATATGT 6979  
4435 ttgcaactgtctccttcaaaagctcgactcaactcgaaggaaggaatgtaccttga 4494  
|||||  
6980 TTGCCACCTGGTCTCTTCAAAAAGCTCGACTTCACCTCCAAAGGAGAGTAAATGCTTGA 7039  
4495 gaactcagtgtaataatccaaaagtggtgtgcaagtgtgacttccagaagaatgaag 4554  
|||||  
7040 GACCTCAGGTGAATTAATCCAAAAGAGTGGCTCAAGTGGACTTCCAGAGACAAATGAAG 7099  
4555 tcaaaagtgtaactactcaaggaagtaaaatctcgtcttaccagcaatgtatgtgaagat 4614  
|||||  
7100 TCACAGGAGTAACTAAGGAGTAATAATCTCTGTTACGACATGATGTGAAGAGAT 7159  
4615 tccatcatctcgaagcaagaatggcaatcaagtgaactccttcttccagaatgtgaag 4674  
|||||  
7160 TCCATCATCTCCAGAGTCAAGATGGCCATCATGTGACTCTTTTTCAGATGCGAAG 7219  
4675 taaagtttttcaaggaaatcaagactccttcaacccgtgtgtgtgaacttctagaaccac 4734  
|||||  
7220 TAAAGGTTTTCAGGGAATCAAGACTCTTCACACCTGTGGTGAACCTCTAGAACCCAC 7279  
4735 cgttactgactgctactcttgaaatcaacccccagaatgtgtgtgaacgaatgtgccgtga 4794  
|||||  
7280 CGTTACTACTCTGCTACCTTCCAAATTCACCCAGAGTTGGGTGACCAAGATTGGCCCTGA 7339  
4795 ggaatgaaggttcgtggcctgagagcacaagaagactctactgacttgaaggaatgttc 4848  
|||||  
7340 GGATGGAGGTTCTGGGCTGCGAGGACAGGACCTCTACTGAGGGTGGCCACTGCG 7393

RESULT 2  
AR110040 4670 bp DNA PAT 14-FEB-2001  
LOCUS AR110040 4670 bp DNA PAT 14-FEB-2001  
DEFINITION Sequence 41 from patent US 6114148.  
ACCESSION AR110040  
VERSION AR110040.1 GI:12826316  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
BASE COUNT 1325 a 1072 c 1036 g 1237 t  
ORIGIN  
Query Match 54.9%; Score 4357.6; DB 9; Length 4670;  
Best Local Similarity 99.8%; Freq. No. 0;  
Matches 4363; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
477 gccaccagaagaatactaccctgggtgcaatgtgactgtcatggaactatataagtgat 536  
|||||  
84 GCCACGAGAAGTACTACTGCTGGGTGACAGTGCATGTCATGAGCTATATCCAAAGTAT 143  
537 cctcgtgtgacgtgctgtggaacgaatctcctcctagaatgtgcaaaatcttccattc 596  
|||||  
144 CTCGGGAGCTCCTGTGAGCAGCAAGATTCTCCCTAAGTGCCTAAATCTTTCCATTC 203  
597 aacacctcgtcgtgtacaaaagaagctcgttgtagaattcacaggttcacatttaac 656  
204 AACACCTCAGTGTGTACAAAAGACTGTTGTGAATTCACGATCACCTTTTCAC 263  
657 atcgcataagcaagggccaccctggaatgtgtctgaagttccatacaccagcgaagtt 716  
264 ATCGCTAACCAAGGCCACCTCGATGGGTGCTGCTAGTCCATCATCCAGCTGAGGTT 323  
717 tatgatataagtgtaacttaagaacatgtgtcccatcctgtatgttcaatgtc 776  
324 TATGATACAGTGTGATTCACCTTAAGAACATGAGCTTCCATCTGTCAGTCTTCATGCT 383  
777 gttgtgtatctcactcgggaagctcttgaaggagtgtaatatgatgcagaagaatca 836  
384 GTTGGGTATCTCTACTGGAAAGCTTCTGAGGAGCTGATATGATGATCAGCCAGTCA 443  
837 aggaagaagaagatgataaagttctccctgtgtgaagccatacatatgtctgcaagtc 896  
444 AGGGAAGAAAGATGATTAAGTCTTCCCTGTGAGACCATATATGTCTGGAGGATC 503  
897 ctgaagaagaatgttccaaatgtgctctgaaccactgtgtcctactactatattctc 956  
504 CTGAAGAAGAAAGTGCATGTGCTGTGACCCACATGCTCTTACTACTCATATCTTCT 563  
957 catgtgacctgtgtaaaagacttgaatcaggcctcatlttgaggccctactagtatga 1016  
564 CATGTGAGCTGTGTAAGAACTTGAATTCAGGCTCATATTGGAGCCCTACTGTATGTGA 623  
1017 gaaggaagctgtgccaagaaagaacacagactctgcacaattatatactacttltgt 1076  
624 GAAAGGAGTCTGGCCAGGAAAGAACACAGACACTTGCACAAATTTATCTACTTTTCT 683  
1077 gtaattgatgaaggaaagtgtgcaactcgaagaacaaagaaactcttgaatgcaagatag 1136  
684 GATTATTGATGAGGAGAAAGTTGGCAGCTCAGAAACAAAGAACTCCTTATGATCAGATAGG 743  
1137 gatgtcgaactgtcctcggcctcgtgcttaaatgtgacacaaagtcagtgtatgtatacagg 1196  
744 GATGCTGATCTGCTCGGCTGGCTTAAATGACACAGTCAATGATTATGTAAACAGG 803

|    |      |   |      |
|----|------|---|------|
| OY | 1197 | tcctcgcagagtcgtatggaatgacaggaatcagttatctgcatgtatggaatg      | 1256 |
| Db | 804  | ttctgcccagagctgattggaatgacaggaatcagttatctgcatgtatggaatg     | 863  |
| OY | 1257 | ggcaccactcctgaaatgcaatcaatctcctgaagtgcaacatcttctgtggaac     | 1316 |
| Db | 864  | ggcaccactcctgaaatgcaatcaatctcctgaagtgcaacatcttctgtggaac     | 923  |
| OY | 1317 | catgcacagagctccttggaatctcgcgaatacttccctaacgctcaaacctctg     | 1376 |
| Db | 924  | catgcacagagctccttggaatctcgcgaatacttccctaacgctcaaacctctg     | 983  |
| OY | 1377 | atggaacttgagcaatcttctactgtttgtcatatctctcccaacaatgtagcatg    | 1436 |
| Db | 984  | atggaacttgagcaatcttctactgtttgtcatatctctcccaacaatgtagcatg    | 1043 |
| OY | 1437 | gaagcttatgtaaaatgtagcagctgtccagaggaaccccaactaagaatgaataat   | 1496 |
| Db | 1044 | gaagcttatgtaaaatgtagcagctgtccagaggaaccccaactaagaatgaataat   | 1103 |
| OY | 1497 | gaagaaggggaagactatgatagtatcttactgtatctgaaatgtagtgcagttt     | 1556 |
| Db | 1104 | gaagaaggggaagactatgatagtatcttactgtatctgaaatgtagtgcagttt     | 1163 |
| OY | 1557 | gatgtatgaactctcctcctcttaccgaatctgcatgctgcaagaagcatctaa      | 1616 |
| Db | 1164 | gatgtatgaactctcctcctcttaccgaatctgcatgctgcaagaagcatctaa      | 1223 |
| OY | 1617 | acttggtatcatctatctgctgtcgaagaaggaagtagtgcacatctctctctc      | 1676 |
| Db | 1224 | acttggtatcatctatctgctgtcgaagaaggaagtagtgcacatctctctctc      | 1283 |
| OY | 1677 | gcccccgatgagaagatgataaagaatctgaacatgagccctcagcgatggt        | 1736 |
| Db | 1284 | gcccccgatgagaagatgataaagaatctgaacatgagccctcagcgatggt        | 1343 |
| OY | 1737 | aggaagtaaaaaaagtcgatttatgcatcacagatgaacacttgaagctgtga       | 1796 |
| Db | 1344 | aggaagtaaaaaaagtcgatttatgcatcacagatgaacacttgaagctgtga       | 1403 |
| OY | 1797 | gctatcagcatgaaatcagaatctctggaaccttacttctatggaagcttgagaca    | 1856 |
| Db | 1404 | gctatcagcatgaaatcagaatctctggaaccttacttctatggaagcttgagaca    | 1463 |
| OY | 1857 | ctgtgtatataatgaagaatcagaacgagacatataatcttaccctcaggaatc      | 1916 |
| Db | 1464 | ctgtgtatataatgaagaatcagaacgagacatataatcttaccctcaggaatc      | 1523 |
| OY | 1917 | actgtatcgcgtctgtatctcaagagatfaccgaagaagtgtaaacatttgaagat    | 1976 |
| Db | 1524 | actgtatcgcgtctgtatctcaagagatfaccgaagaagtgtaaacatttgaagat    | 1583 |
| OY | 1977 | tttccaattctgcagaggaataltcaaatataaagtgaacagtgtcgtgaagaatg    | 2036 |
| Db | 1584 | tttccaattctgcagaggaataltcaaatataaagtgaacagtgtcgtgaagaatg    | 1643 |
| OY | 2037 | ccaactaatcagatcctcgtgtgctgtgacccgtatctactagtcttctgttaagt    | 2096 |
| Db | 1644 | ccaactaatcagatcctcgtgtgctgtgacccgtatctactagtcttctgttaagt    | 1703 |
| OY | 2097 | agaagctcagctcagagatcattgagccctcctcactcgtgtaagaagaatctgag    | 2156 |
| Db | 1704 | agaagctcagctcagagatcattgagccctcctcactcgtgtaagaagaatctgag    | 1763 |
| OY | 2157 | caaagaagaaacagataatgtcagacaagaagaaatgtcatctctgttctgtatg     | 2216 |
| Db | 1764 | caaagaagaaacagataatgtcagacaagaagaaatgtcatctctgttctgtatg     | 1823 |
| OY | 2217 | gagaacccgaagctgtgtactcacaagaataatacagaagcttctcccaatccagctga | 2276 |
| Db | 1824 | gagaacccgaagctgtgtactcacaagaataatacagaagcttctcccaatccagctga | 1883 |

---

|    |      |   |      |
|----|------|---|------|
| OY | 2277 | gtcagactgaagatccagagatcccaagctccccaatcatgtcacagatcaatgtcat    | 2336 |
| Db | 1884 | gtcagactgaagatccagagatcccaagctccccaatcatgtcacagatcaatgtcat    | 1943 |
| OY | 2337 | gttttgaatgttgcagatgtgcagttgttctgcatgaggtgcatatgtatctcta       | 2396 |
| Db | 1944 | gttttgaatgttgcagatgtgcagttgttctgcatgaggtgcatatgtatctcta       | 2003 |
| OY | 2397 | agcattgagacagacatcactctctctgtctctctctctctctctctctctctct       | 2456 |
| Db | 2004 | agcattgagacagacatcactctctctgtctctctctctctctctctctctctct       | 2063 |
| OY | 2457 | aaaatgctatgaagacacacatccatccatctccatctcagaggaactgtctcatg      | 2516 |
| Db | 2064 | aaaatgctatgaagacacacatccatccatctccatctcagaggaactgtctcatg      | 2123 |
| OY | 2517 | tcgatggaacccaggtctatgatatctgggtgtccacaactcagactctggaacaga     | 2576 |
| Db | 2124 | tcgatggaacccaggtctatgatatctgggtgtccacaactcagactctggaacaga     | 2183 |
| OY | 2577 | ggcatgacgccttactgaagttctagtttgacaagaacactggtatattatagag       | 2636 |
| Db | 2184 | ggcatgacgccttactgaagttctagtttgacaagaacactggtatattatagag       | 2243 |
| OY | 2637 | gacagtatgaagatattcaagcatctgtcgtgagtaaaacaatgtccatgaacaga      | 2696 |
| Db | 2244 | gacagtatgaagatattcaagcatctgtcgtgagtaaaacaatgtccatgaacaga      | 2303 |
| OY | 2697 | agcttctccgaatctcaagacacactgacatgagcaaaagaatlttaatgcccacca     | 2756 |
| Db | 2304 | agcttctccgaatctcaagacacactgacatgagcaaaagaatlttaatgcccacca     | 2363 |
| OY | 2757 | ccagctctgaaacgcacatcaaacgggaataaactgttactactctctcagctcaatgaag | 2816 |
| Db | 2364 | ccagctctgaaacgcacatcaaacgggaataaactgttactactctctcagctcaatgaag | 2423 |
| OY | 2817 | gaacttgactatgatacacaatcactgagtgaaatgaagaagaagatttacaattat     | 2876 |
| Db | 2424 | gaacttgactatgatacacaatcactgagtgaaatgaagaagaagatttacaattat     | 2483 |
| OY | 2877 | gatgagatgaaatcagaagcccccagcttcaaaagaacacagacatattatct         | 2936 |
| Db | 2484 | gatgagatgaaatcagaagcccccagcttcaaaagaacacagacatattatct         | 2543 |
| OY | 2937 | gctgcagtggaagctctggatataatgagatgagtagtcccccaatgttctgaagaac    | 2996 |
| Db | 2544 | gctgcagtggaagctctggatataatgagatgagtagtcccccaatgttctgaagaac    | 2603 |
| OY | 2997 | agggctcagatgagcagtgctccctcagttcaagaagaagtgttctcagaatctatg     | 3056 |
| Db | 2604 | agggctcagatgagcagtgctccctcagttcaagaagaagtgttctcagaatctatg     | 2663 |
| OY | 3057 | ggctccttactcagcccttataccgtggaagaaactaaatgaacatttggagcctg      | 3116 |
| Db | 2664 | ggctccttactcagcccttataccgtggaagaaactaaatgaacatttggagcctg      | 2723 |
| OY | 3117 | ccatatagaagcagaagttgaaagataatacatcgtgtacacttcaagaatacagccct   | 3176 |
| Db | 2724 | ccatatagaagcagaagttgaaagataatacatcgtgtacacttcaagaatacagccct   | 2783 |
| OY | 3177 | gctccctatctctctatctcagcttattcttctgaagaaatcagaagcagaagaga      | 3236 |
| Db | 2784 | gctccctatctctctatctcagcttattcttctgaagaaatcagaagcagaagaga      | 2843 |
| OY | 3237 | gaacctagaanaaactgtcgaagcttaatgaacacaaacttacttltgaaagtgc       | 3296 |
| Db | 2844 | gaacctagaanaaactgtcgaagcttaatgaacacaaacttacttltgaaagtgc       | 2903 |
| OY | 3297 | catcatatggaacccacacaaagatgagtttgagctgcaagcctggtctatctctg      | 3356 |
| Db | 2904 | catcatatggaacccacacaaagatgagtttgagctgcaagcctggtctatctctg      | 2963 |
| OY | 3357 | gttgacctggaanaagatgtgacatcagccctgatttgacccctctgtgtcgcacac     | 3416 |

```

|||||
Db 2964 GTTGGACCTGGAAAAAGATGTGCACTCAGGCCGTGATTGACCCCTTGTGCTGCCACACT 3023
3417 aacaacacgaaccctgcctcattggagaagaagtgcagatcagaatctgctgttttc 3476
|||||
Db 3024 AACACACTGAACCCCTGCTCATGTGGAGACAAAGTACAGTACAGAAATTTGCTGTTTTTC 3083
3477 accacctcttgatgagaccgaagaactggtactcactgtaaaataatggaaagaactgcag 3536
|||||
Db 3084 ACCATCTTTGATGAGAACCCAAAAGCTGTACTTCACTGAAATATGAAAGAAACCTCAG 3143
3537 gctccctgcataatccagatggagaatcccacttttaagaagaattatcgcttcattgca 3596
|||||
Db 3144 GCTCCCTCAATATCCAGATGGAAGATCCACTTTTAAAGAGAAATATGCTTCCATGCA 3203
3597 atcaatgctacataatgatactactaccctggtcttaagtaatgctcagaagataaagatt 3656
|||||
Db 3204 ATCAATGGCTACATATATGATACACTACTGCTTATGATGCTCAGATCAAAAGGATT 3263
3657 cgatgtaattcgtcagcatgggcagcaatgaataacatcattctatctcattcattgga 3716
|||||
Db 3264 CGATGATATGCTCAGCATGGGCAAGCAATGAACATCATTTCTATTCATTTCATTCAGTGA 3323
3717 catgtgtcactgtatcgaaaaaagaagaataaaatggcactgtacaaatctctatcca 3776
|||||
Db 3324 CATGTGTTCACTGTACGAAAAAAGAGAGATTAATAATGCACTGACATCTATTCACA 3383
3777 ggtgttttgaagacagtgaanaatttaacatcgaagcgtggaatttgggggttggaatgc 3836
|||||
Db 3384 GGTGTTTTTGAACAGTGAATGTTACCATTCMAAGCTGGAATTTGGGGGTGGAAATGC 3443
3837 ctatctggcagacatctacatgcttggaatgacacatttcttggctgacagcaataag 3896
|||||
Db 3444 CTATTTGGCGAGCATCTACATCTGGAGATGACACACTTTTCTGGTGTACAGCAATAG 3503
3897 tgtcgaactcccttggaatggtctctgacacatlaagaaatttcaagtttcaagttca 3956
|||||
Db 3504 TGTCAGACTCCCTGGGATGGCTTCTGACACATTGAGATTTCAGATTTCACAGCTTCA 3563
3957 ggaacatctgagcaatggggcccaagcgtgcagacatctatctcgaatcaaatcaat 4016
|||||
Db 3564 GGACAAATATGACAGTGGGCCCAAGCTGGCCAGACCTTCATTATTTCCGATCAATCAAT 3623
4017 gctctggagcaccgaagacccttctcttgatcaagtgatctgttggcaccgaatgalt 4076
|||||
Db 3624 GCCTGGACACCAAGAGGCCCTTTCTTGATCAAGTGTGATCTGTTGGCACCAATGATT 3683
4077 attcaagcatcaagaaccgaaggtgcctgcaagaatttccaagctctcaactctcag 4136
|||||
Db 3684 ATTCACGGCATCAAGACCCAGGGTCCCGTCAGAAAGTTCTCCAGCCTTACATCTCTCAG 3743
4137 tttaatcatcatgatagtcttgaatggagaagtgcaagactatcgaagaatctcact 4196
|||||
Db 3744 TTTATCATCATGTTATAGTCTTATGGAAGAGAGTGGACAGACTTATTCAGGAATTTCCACT 3803
4197 ggaaccttaatgyltctcttggaatgltgaa ttcaatcctggaataaaacacatalttct 4256
|||||
Db 3804 GGAACCTTAATGCTCTCTTTGGCAATGGAATCATCTGGGATTAATAACATATATTTT 3863
4257 aaacctccaattatctgcgatataatcgttctgcaaccaactcatatagcattcgagc 4316
|||||
Db 3864 AACCTCAATATATGCTGATACATCGTTTGCACCCCAACTTATATGATATTCACAC 3923
4317 actctgcagaggttgatgggtgctgtaattaaatgattgagaatgcatctggaaatgc 4376
|||||
Db 3924 ACTCTTCCAGATGAGTTATGGGCTGTGATTTAAATAGTTGACGATGCCATTGGGAATG 3983
4377 gagagtaagaacaataatcagatgcacagatactactgtcacttcaacaataatgltc 4436
|||||
Db 3984 GAGAGTAAGCAATATCATGATGCACAGATTACTGCTTCACTACTTTCACATATATGTTT 4043
4437 ggcacctggtctcctctcaaaagctcgaactcgaactccaaaggaggaataatgctcggaga 4496
|||||

```

```

Db 4044 GCCACCTGCTCTCTTCAAAAGCTGCACCTTCACTCCAAAGGAGGATATGCTTGAGA 4103
4497 cctcaagtgtaataatccaaagaagtgctgcgaatggaacttccaaagaacatgaatc 4556
|||||
Db 4104 CCTCAGCTCAATATATCCAAAGAGTGCTGCAAGTGGACTTCCAAACATATGAAATC 4163
4557 acagagaataactcgaaggaglaaataatctgcttaccagaatgaltgtaaggatctc 4616
|||||
Db 4164 ACAGAGTAATACATCAGAGAGTAATAAATCTGTTACAGCATGTATGTGAAGAGATTTC 4223
4617 ctcatctcagaagatcaagaatgcatcgaatggaactctcttttccaagaatggcaagta 4676
|||||
Db 4224 CTATCTCCAGAGTCAATGAATGATGATGATGATGATGATGATGATGATGATGATGATG 4283
4677 aaggttttcaagggaataaagaatcccttcaacactgtggtgaactctctagaccacg 4736
|||||
Db 4284 AAGGTTTTTCAGGGAATCAACATCTTACACACTGTGTGTAATCTCTAGACCCACCG 4343
4737 ttaactgactgctactccttgaatcaccaccagaatggtggtgacagaatggtcctgag 4796
|||||
Db 4344 TTACTGACTGCTACCTTGAATTCACCCCGAGAGTGGTGCACAGATTGCCCTGAG 4403
4797 atggaggtctggtgctgcgagagcagaagacactctactgactclogagcagttc 4848
|||||
Db 4404 ATGAGGTTCTGTGGCTGCGAGGACAGACCTCTACTGAGGGGTGCGCACTGC 4455

RESULT 3
108644
LOCUS 108644 4278 bp PAT 02-DEC-1994
DEFINITION Sequence 4 from Patent WO 8800831.
ACCESSION 108644
VERSION 108644.1 GI:588650
KEYWORDS
SOURCE
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 4278)
AUTHORS
Pasek, M.P.
TITLE
DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR
VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE
POLYPEPTIDES IN HIGH YIELDS
JOURNAL
Patent: WO 8800831-A 4 11-FEB-1988;
FEATURES
Location/Qualifiers
source 1..4278
BASE COUNT 1244 a 942 c 946 g 1146 t
ORIGIN
Query Match 52.2%; Score 4148.8; DB 10; Length 4278;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 4270; Conservative 0; Mismatches 2; Indels 90; Gaps 3;

```

```
OY 777 gttagtgcatacttggaagctctgaaggagctgaataatgatgcagaccagtica 836
|||||
Db 301 GTTGGTGTATCTACTGGAAGAGCTTCTAGGGAGCTGATATGATGATCAGACCACTGTA 360
OY 837 agggagaaagaaatgatataaagctctcccttggtggaagccacacacatgctcgaagtc 896
|||||
Db 361 AGGGAGAAAGAAAGATGATAAAGTCTTCCCTGGTGGAGCCATACATATGTCGGAGGTC 420
OY 897 ctgaaaggaatggtcacaatggtcctcgaacccatgctcctactactacatcttctc 956
|||||
Db 421 CTGAAGAGAAATGTCATGCTGACCTGACCCACAGTGTCTTACTACTCATATCTTCT 480
OY 957 catgtggagctcggtaaaagactgaatccaggcccatctggagccctactgaatgata 1016
|||||
Db 481 CATGGAGACTGGCCAAAGAAAGACACACACTTGACAAATTTATACTTCTTCT 540
OY 1017 gaaggagctcggcgaagaaagacacagacactgacacaaatataactacttctgct 1076
|||||
Db 541 GAAGGAGCTGGCCAAAGAAAGACACACACTTGACAAATTTATACTTCTTCT 600
OY 1077 gtattgtatgaaggaagaaagtgtgcacccagaaacaaagaaactccttgaatgcagatagg 1136
|||||
Db 601 GTATTGTGATGAAGGAAAAAGTTGGCAGTCAGAAACAAAGAACTCCTTGATGACAGATAG 660
OY 1137 gatgtcatctgtcgcggcctggcctaaatgacacagtcacatggttatgtaaacagg 1196
|||||
Db 661 GATGCTGATCTGCTCGGGCTGGGCTTAATAATGCACACAGTCATGTTATGTAACAGG 720
OY 1197 tctctg---ccaagttctgattgtaatgacacagaaatcagtcatactgacatgta 1253
|||||
Db 721 TCTCTGCTACCAAGGCTGATTGGATGCCAGGAATCAGTCTATGGCATGTGATGGA 780
OY 1254 atgggacacactcctgaagatgcacatcaatctcctgaaggtcacacatctcctgtgagg 1313
|||||
Db 781 ATGGCACACCTCTCGAAGTGCACATATCTCCGAAGGTCAACACTTCTGTGAGG 840
OY 1314 aacctgcagagcgctccttggaatctcgcaataactctcctactgctcacaacatc 1373
|||||
Db 841 AACCTGCCAGGCGCTCTTGGAAATCTGCGCAATTACTTCTTACTGCTCAAAACATC 900
OY 1374 ttgatgacacttggaacagttctactglttctgacatactctccacacaacatgagtc 1433
|||||
Db 901 TTGATGAGACTTGGACAGTCTTCTACTGTTTGTCTATCTCTCCACCAACATGATGGC 960
OY 1434 atggaagcttatgtcacaagtgaacagcgtctcagaaggaacccccacatcgaaatga 1493
|||||
Db 961 ATGGAAGCTTATGTCAAAGTAGACAGCTGTCCAGAGAAACCCCAACTAGATGAAAAAT 1020
OY 1494 aatgaagaagcggaaagactatgatgatctactgaatctgaaatgagatgagtcagg 1553
|||||
Db 1021 AATGAAGAAGCGGAAGACTATGATGATGATCTTACTGATTTCTGAAATGATGATGTCAG 1080
OY 1554 ttgatgatacacaactcctccttcccttcaacaaatcgctcagttgcccagaagacatcct 1613
|||||
Db 1081 TTTGATGATGACAACCTCTCTTCTTATCCAAATTCGCTCAGTTGCCAAGAAGATCTCT 1140
OY 1614 aaacttgggttaacttaacttgcgtcgtgaaggaaggaacatgagactgctcctttagtc 1673
|||||
Db 1141 AAAAAGCTGGGTACATTTATGCTGCTGAAAGAGAGACTGGACTATGCTCCCTTAATC 1200
OY 1674 ctgcgccccgatacagaagttataaagtaacatattggaacaaatgccccacagagatt 1733
|||||
Db 1201 CTCGCCCGCATGACAGAGATTTATTAAGTCAATTTTGAACAAATGCGCCCTCAGCGGATT 1260
OY 1734 ggtggaagatacaaaaaagtcogatttaatggtcacacagatgaacacatcttaagactcgt 1793
|||||
Db 1261 GGTAGGAGATGCAAAAAAGTCCGATTATGGCATACAGATGAACCTTTAAGCTCGT 1320
OY 1794 gaagctatcagcagaaatcaggaatccttggaccttacttattgggaaggtggagac 1853
|||||
Db 1321 GAAGCTATTACAGCATGATGATGAGAACTTGGAGCTTACTTTATGGGAGATGGAGAC 1380
OY 1854 aacctgttattatatttaagaatcgaagcaagcagaccataataacatctaccacaga 1913
|||||
Db 1381 ACACGTGTTGATTATTTAAGATTCACAGACAGACATATACATCTTACCCTCAGCGA 1440
OY 1914 atcaactgatactcgtcccttggatlatcaaggaagatcacaaaaggtgtaaaaactltgaag 1973
|||||
Db 1441 ATCACTGATGTCGCGCTTGTATTCAGAGAGATTACCAAAAAGGTGAACATTTGAAG 1500
OY 1974 gatttccaattctgcagaggaagaatatcaataataaaygaacagtgatgagaagat 2033
|||||
Db 1501 GATTTTCCAAATTCGCCGAGGAATATTTCAAAATATAAATGACAGTGCATGTGAAAT 1560
OY 2034 gggccaactaaataagatccctcgggtgcctggaacccgatactactagttcgtaaatg 2093
|||||
Db 1561 GGGCACTAATATCAGATCTCGGCGCTGACCCGCTATTACTAGTTGGTTAATATG 1620
OY 2094 gagaagatactagcttcaagacataatgagccctccatcctcgtacataaagaatctgta 2153
|||||
Db 1621 GAGAGAGATCTAGCTTCAGAGCTCATTTGCCCTCTCCATCTGCTACAAAGAACTGTGA 1680
OY 2154 gatcaagaagaaacagagatatgttcagacaaggaatgtaacatcgttcttctgtaatt 2213
|||||
Db 1681 GATCAAGAGGAAGAACCATATATGTACAGACAAGAGAAATGTATCTGTTTCTGTATTT 1740
OY 2214 gatgagaacgaagctggttacctcaagagaatatacaagcttctcccaatccagct 2273
|||||
Db 1741 GATGAGAACCCAGCTGCTACTCTCAGAGAGATATACAGCCTTTCTCCCAATCCAGCT 1800
OY 2274 ggaatgcagcttgaagatccagagttccaaagcctccaaatcatatgacagatcaatgctc 2333
|||||
Db 1801 GGAGTGCAGCTTGAGAGATCCAGAGATTCACAGCTCCCAACATCATGACAGCATCAATGGC 1860
OY 2334 tatgttttgaatgttggcaatgtgcagttglttgttgcaatgggtggtgacatgtaatt 2393
|||||
Db 1861 TATGTTTGTGATGATGTCAGTGTGTCACTTGTGTTGATGAGGTGACATCTGTATCT 1920
OY 2394 ctaagcatgtgacacagaacagactcctctctctctcctcctctctgataccttcaaa 2453
|||||
Db 1921 CTAAGCATTTGAGACAGACAGACTCTCTTCTGTGCTTCTCTGATATACCTTCAAA 1980
OY 2454 cacaanaatgtaataaagacacacacacacatccatctccatctcaaggaacatgctctc 2513
|||||
Db 1981 CACAAAATGCTATGAAAGACACACTCCACTATTTCCATCTCAGGAGAAACGTCTTC 2040
OY 2514 atgtcgaatggaanaacccaagctctatgtaattctgggtggtgcacaactcgaacttccgaac 2573
|||||
Db 2041 ATGTCGATGGAAGAACCCAGGCTATGATGATCTGAGGTCACCAACTCAGACTTTCGGAAC 2100
OY 2574 agagcaatgacccgcttactgaagttctagttgtgacaagaacacatggtatattac 2633
|||||
Db 2101 AGAGCATGACCGCTTACTGAAGTTCTGATGTTGTGACAAAGAACACTGGATTTATTAC 2160
OY 2634 gaggacagttatgaagatattcagcatcttgctgtgataaanaaatgccaatgtaacca 2693
|||||
Db 2161 GAGACAGTTATGAAGATATTTCAGCATACTTCTGATGAATAAAACAATGCACTTGAACCA 2220
OY 2694 agaagcttctcccaagaattcaagacacccctagcactgagcaaaagcaatlaatgccaac 2733
|||||
Db 2221 AGA----- 2223
OY 2754 ccaccagctctgaaacgcatacgaaggaataaactcgtactactctcagtcagatcaa 2813
|||||
Db 2224 -----GAATTAACCTGTAATCTTACTCTTCACTGATCAATCA 2256
OY 2814 gaggaattgactatgataatcacatatcagttgaatgaaaggaagaatttgacatt 2873
|||||
Db 2257 GAGGAAATTTGACTATGATGATACATATCACTGTAATGAAAGAGGAAGATTTGACATT 2316
OY 2874 tatgatgagatgaaataatcagcccccgagccttcaaaaagaaacagacatatctt 2933
|||||
Db 2317 TATGATGAGATGAAATATCAGAGCCCCGAGCTTTTCAAAAGAAACGACACTATTTT 2376
OY 2934 atgtcgtcagtgagagagcctcgtgattatgataatgatactcccaatglttctaaga 2993
```



|||||  
Db 2377 ATTGCTGAGTGGAGAGGCTTCGGATTATGGATGAGTAGCTCCCAATGTTCTAATA 2436  
Oy 2994 aacaaggctcagagtgagtgctccctcagtcacaagaattgttttcaggaattact 3053  
Db 2437 AACAGGGCTCAGAGTGGCAGTCCCTCAGTTCAGAAAGTTGTTTCCAGGAATTTACT 2496  
Oy 3054 gatgctccttactcaagccctataccgttgagaaactaatagaatttggagctcctg 3113  
Db 2497 GATGGCTCTTACTACAGCCCTTATACGCTGGAGAACTAATGAACATTTGGACTCTG 2556  
Oy 3114 gggccataataaagcagaagtgtgaagataatcatatgttaacttcgaagaatcagcc 3173  
Db 2557 GGGCCATATATAAGAGCAGAAAGTTGAAGATATCATGTGTAACCTTTCAGAAATCAGGCC 2616  
Oy 3174 tctcgtccctatctccttactatctagcccttattcttaagaaagatcagaagcaga 3233  
Db 2617 TCTGCTCCCTATCTCTTCTATCTAGCCCTTATTCTTATGAGAGAAATCAGAGCAGAGA 2676  
Oy 3234 gcagaacctagaanaaacttgttcaagcctaaagaaacaaacttacttctggaagtgtg 3293  
Db 2677 GCAGAACTAGAAAACCTTGTCAAGCCTAATGAACCAAAACCTTACTTTGGAAGTG 2736  
Oy 3294 caaatcatataggcccaactaaagaatgagttgactgcaaaagcctgagcttattctct 3353  
Db 2737 CAACATCATATAGGACACCACTAAGATGAGTTGACTGCAAGCCTGGGCTTATTTCTCT 2796  
Oy 3354 gatcttgaccttgaaaaagatgtgcactcagagccttgatctggaacccctctgtctcag 3413  
Db 2797 GATGTTGACCTGGAAAAGATGTGCACCTAGGCTGATTTGGACCCTTGTGCTGCCAC 2856  
Oy 3414 actaacacacatgaaccctgtcctcaatggagaaagatgaacgaatctgtctgtt 3473  
Db 2857 ACTAACACACACTGAACCCCTGCTCATGTGGAGACAAAGTACAGTACAGAAATTTCTGTTT 2916  
Oy 3474 c---tcacatctttagatgagaccaaagctgtgtacttcaactgaaatctggaagaac 3530  
Db 2917 TTCCCTACACTTGTGATGAGACCAAAAGCTGTACTTACGAAAATVTGAAAACAAAC 2976  
Oy 3531 tgcagagctccctcgaataatccagaatgagaatcccaactttaaagaatatacgtctc 3590  
Db 2977 TGCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAGATTTATCCCTTC 3036  
Oy 3591 catgcaatcaatggtcactataatggaatacactcgtgctgtgaatggcctcagaatcaa 3650  
Db 3037 CATGCAATCAATGGCTACATATGATGATCACTACCTGGCTTGTATGGCTCAGGATCAA 3096  
Oy 3651 aggaattgattgattctcctcaggaatggagcaatgaaacatccattcttcaatttc 3710  
Db 3097 AGGATTCAGATGATCTCTGTCAGCATGGGACCAATGAAACATTCATTCATTTTC 3156  
Oy 3711 agtggacaatgttctcactgtaacgaanaaagaagagatataaaatggtcactgtacaatcc 3770  
Db 3157 AGTGAGACATGTGTTCTACGTACGAAAAGAGAGATATAAAATGGCACTGTACAAATCTC 3216  
Oy 3771 tatccaggtgttttttgagaacagtgtgaatgttaccatccaaagctgtgaatttggcgggtg 3830  
Db 3217 TATCCAGCTGTTTGTGAACAGTGAATGTATACATCCAAAGCTGGAAATTTGGCGGGTG 3276  
Oy 3831 gaatgacctatttggagacatctacatgcttgatgagacacatttctgtgtacacg 3890  
Db 3277 GAATGCCCTTATTTGGGACATCTTACATGCTGGATGAGACACTTTTCTGTGTACACG 3336  
Oy 3891 aataagtgtagaacctcccttggaatgtgtcttgacacatctagaatctcagatttcaatata 3950  
Db 3337 AATAAGTGTACAGACTCCCTCGGATGGTCTTGACACATATGAGATTTTCAGATTATA 3396  
Oy 3951 gctcagaacaatatgacagtgggcccaaaagctgtgccagactcatatctcagatca 4010  
Db 3397 GCTTCAGAGACATATAGCAGTGGGCCCCAAAGCTGGCCAGACTTCATTAATTCGGATCA 3456  
Oy 4011 atcaatgcttgagagacaaagagcccttcttgatcaagtgatgtgtgttgaccca 4070  
|||||

Db 3457 ATCAATGCTGAGACACCAAGAGCCCTTTTCTTGATCAAGTGATCTGTTGGCACCA 3516  
Oy 4071 atgattatcacgacatcaagaaccaagggtgccgttcaagaattctccagcctatactc 4130  
Db 3517 ATGATTATTCAGGGCATCAAGACCAGGCTGCCCTCAGAAATTTTCCAGGCTCTACATC 3576  
Oy 4131 tctcagttatcaatcatatagcttgaatggaaagtggtcagaacttatacagaat 4190  
Db 3577 TCTCATTTATCATGATGATGATGCTTTCATGGAAAGTGGCAGACATTAATCGAGGAAT 3636  
Oy 4191 tccagtgaaacttaatggtcttcttgcaatgtggaattatctcgtggaataaacacat 4250  
Db 3637 TCCACTGGAACTTAAATGGCTTTTGGCAATGTGATTCATCTGGGATTAACACCAAT 3696  
Oy 4251 attttaaccctcaatcttctcgtatataccgttgcgaaccaaccatatagaatt 4310  
Db 3697 ATTTTAACTCCATATTTCTCGATTCATCCGTTTGGACCAACATTAATATGATTT 3756  
Oy 4311 gcgaacactctcgaatgagttgagtggtgtgatttaaatagttgacagatgccatg 4370  
Db 3757 CCGAGCACTCTTCGATGAGATGGCTGTGATTTAAATAGTTCAGCATGCCATTTG 3816  
Oy 4371 ggaatgagaatfaaagcaatatcagatgacagatctactgcttcaacttacaat 4430  
Db 3817 GGAATGGAAGTAAACCATATACATGACATGACAGATTCTGCTTACCTTATACCAAT 3876  
Oy 4431 atgttgcacactgtgtctccttcaaaagctcgaacttcaactccaagggagagtaatgcc 4490  
Db 3877 ATGTTTGGCACCTGGTCTCTTCAAAAGCTGACTTCACTCAAGGAGAGATTAATGCC 3936  
Oy 4491 tggagaactcgaatgaaataatccaaagaatgtgtcgaagtgtgacttccagaagaatg 4550  
Db 3937 TGGAGACCTCAGATGAATATCAAAAGAGTGGCTGCAAGGTGACTTCCAAAGCAATG 3996  
Oy 4551 aaagtcaagagtgtaactactcgaagagtaaaatctcgttacaagaatglatgtgaag 4610  
Db 3997 AAAGTACAGAGTAACTACTGAGGAGTAAATCTGCTTACGAGATTAATGTGAAG 4056  
Oy 4611 gatttccatctcagcagatcaagaatgccaatgagtcagtcacttcttcttcaagaatggc 4670  
Db 4057 GAGTTCCTCATCTCCAGGACGATCAAGATGGCCATCAGTGGAGCTCTTTTCCAGATGGC 4116  
Oy 4671 aaagttaaagtttccaaggaaatcaagaactcctcaacactgtgtgtaactcttagac 4730  
Db 4117 AAAGTAAAGTTTTCAGGAAATCAAGACTCTTCAACACTGTGTGAATCTCTTACAC 4176  
Oy 4731 ccaacgttaactgaactcgaactcgaattcaaccccaaggttgggttccacagattgcc 4790  
Db 4177 CCACGCTTACTACTGCTACTCTTGAATTCACCCCAAGATTGGGTGCACCAATGCC 4236  
Oy 4791 ctgagatggaagttctgtgctgcagagcacagaacctctac 4832  
Db 4237 CTGAGGATGAGGTTCTGTGGCTGCGAGGACAGACACTCTAC 4278  
|||||

RESULT 4  
LOCUS 108643 4281 bp  
DEFINITION Sequence 3 from Patent WO 8800831. PAT 02-DEC-1994  
ACCESSION 108643  
VERSION 108643.1 GI:588649  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4281)  
AUTHORS Pasek,M.P.  
TITLE DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE  
JOURNAL POLYPEPTIDES IN HIGH YIELDS  
FEATURES Patent: WO 8800831-A 3 11-FEB-1988;  
SOURCE location/Qualifiers  
1. 4281

BASE COUNT 1245 a 943 c 946 g 1147 t  
ORIGIN

Query Match 52.28; Score 4147.2; DB 10; Length 4281;  
Best Local Similarity 97.9%; Pred. No. 0;  
Matches 4269; Conservative 0; Mismatches 3; Indels 90; Gaps 3;

QY 477 gccaccgaagaatactactcgtggtcagtggaactgtcatgtggaactatataatgcaagatgat 536  
|||||  
Db 4 gccaccgaagaatactactcgtggtcagtggaactgtcatgtggaactatataatgcaagatgat 63  
QY 537 ctgcgtgagcgtcctgtggaacgaagatttccctctagagtgccaaaatcttccattc 596  
|||||  
Db 64 ctgcgtgagcgtcctgtggaacgaagatttccctctagagtgccaaaatcttccattc 123  
QY 597 aacacccatgctcgttcaaaaaaagactctgttggtagaatcaaggttcacattcacaac 656  
|||||  
Db 124 AACACCTCAGTCTGCTACAAAAGACTCTGTTGTAGAAATTCACGAGATCACCTTTTCAAC 183  
QY 657 atgcctaaagcgaagccacccctgtgagtggtctgctgaagtcaccatccaggtgaagtt 716  
|||||  
Db 184 ATCCCTAAGCCCAAGGCCACCTGGATGGGTCTGCTAGGTCTCTACCATCCAGGGTGAAGGTT 243  
QY 717 tatgatacaagtgtcattactaactaaagaacatggtccctccatcctctcagctctcatgct 776  
|||||  
Db 244 TATGATACAGTGGTCTATTACTCTTAAGAACTAGTGGCTCCCACTCTGTCAGTCTTCATGCT 303  
QY 777 gttgtgtatcctactctgaaagactctctgaggagctgaataatgaatgaacagaccatcaa 836  
|||||  
Db 304 GTTGCTGATCTCTACTGAAAGCTTCTGAGGAGCTGATATGATGATCAGACAGTCAA 363  
QY 837 agggagaagaagatgtataaagcttccctggtgtggaagccaataatgtctggaagttc 896  
|||||  
Db 364 AGGAGAGAAAGATGATTAAGTCTCCCTGGTGAAGCCATACATATGCTGTGACAGTCC 423  
QY 897 ctgaaagaagaatgttccaaatggtcctctgacccactgtgcttactactactactacttct 956  
|||||  
Db 424 CTGAAAGAGATGCTCCAAATGGCTCTGACCCACTGTGCTCTACTCTCTATCTTTCT 483  
QY 957 catgtgacctgtgtaaaagactgaattcaggccctcatctgagccctactagtaatgta 1016  
|||||  
Db 484 CATGTGACCTGTGTAAGAAAGCTTAATTCAGGCTCATTTGGAGCCCTACTAGTATGTA 543  
QY 1017 gaaggaggtctggtccaaaggaagaacacagacctgtgcaaaaatttaactacttctgct 1076  
|||||  
Db 544 GAAAGGAGTGTGGCCAGGAAAGACACACACCTTGACAAATTTATACACTTTTGTCT 603  
QY 1077 gtaattgaatgaagaagaattgtgcaactcagaagaacagacctctgtatgacagatagg 1136  
|||||  
Db 604 GTATTATATGAAGGAAAGTTGGACTCTCAGAAACAAAGAACTCTTGATGACAGGATAGG 663  
QY 1137 gatgtcgtacatctgtcgtggtcgtgcttaaaatgacacagtcacatggttlaatgtaaacag 1196  
|||||  
Db 664 GATGCTCATCTGCTGGGCTGGGCTGAAATGCAACACAGTCAATGGTTATGTAACAGG 723  
QY 1197 tctctg----cgaagtcgtatgtgatacgaagaatcagtcataatgtgacatgtatgga 1253  
|||||  
Db 724 TCTCTGTACCAAGTCTGATTTGATGTCACAGGAAATCAGTCTATTGGAATGTTGGA 783  
QY 1254 atgggaacacactcctgaagtcactaataatctctggaaggtcacaattctctgttaag 1313  
|||||  
Db 784 ATGGGACACACTCTCTGAAGTGCATCAATATCTCTGAAAGTGCACAAATTTCTGTAGG 843  
QY 1314 aaccatgcgaagcgtctctgtgaaatctgcocaaatgaacttctctactgtcacaacatc 1373  
|||||  
Db 844 AACCATGCGCAGGCGCTCTGGAATCTCGCAATTAACCTTCTTACTGTGCAAAACATCTC 903  
QY 1374 ttgatgacctgtgacaggttctactgttctgtgcatatctcttcccaacaatgatgagc 1433  
|||||  
Db 904 TTGATGAGACCTTGACAGTTTCTACTGTTTGTTCATATCTCTCCACCAACATGATGGC 963

QY 1434 atggaagcttatgtcaaaagtacagacagctgtccagaggaagcccaactacgaatgaaaaat 1493  
|||||  
Db 964 ATGGAAGCTTATGTCAAAAGTAGACAGCTGTCCAGAGGAAACCCCACTCGAATGAAAAAT 1023  
QY 1494 aatgaagaagcgggaagactatgatgatgatactactgattctgaaatgagatgtgtcagg 1553  
|||||  
Db 1024 AATGAAGAGCGGAAGACTATGATGATGATGATCTACTGATCTGAAATGATGATGTGTCAGG 1083  
QY 1554 ttgatgatgacactctccctctccttcttcaaatctgactagttgtccaaagaacatcct 1613  
|||||  
Db 1084 TTTGATGATGACAACTCTCTCTCTTATTCAAAATTCGCTCAGTTGCTCCAAAGAACATCTCT 1143  
QY 1614 aaaactgtggtacattacatgtcgtcgaagaaggaagcgtggactatgtcccttactc 1673  
|||||  
Db 1144 AAAACTTGCGTACATTTACTTGTCTGTAAGAGAGAGACTGGACTATGCTCCCTTACTC 1203  
QY 1674 ctgcgcccccgaatgacagaagttataaaagtcaataatltgaaacaatlygcccctcaagat 1733  
|||||  
Db 1204 CTGCCCCCGATGACAGAGATTATAAAGTCAATATTGAAACAATGGCCCTCAGCGGATTT 1263  
QY 1734 ggttagaagttacaaaaaagtcagattatgtgcatacacagatgaaccttgaagctcgt 1793  
|||||  
Db 1264 GGTAGGAAGTCAAAAAAGTCCGATTTATGGCATACACAGATGAACCTTTAAGACTCGT 1323  
QY 1794 gaagctatcagacatgaatcagaagaatcttggaaccttacttcttgagggaagttggaac 1853  
|||||  
Db 1324 GAAGCTATTGACATGATGATGAGAACTTTGGACCTTACTTATGGGGAAGTTGGAGAC 1383  
QY 1854 acaactgtgatatatatltaaagaatcaagcaagcagacacataataactatccctcaagca 1913  
|||||  
Db 1384 ACACGTGTTGATTATATTAAACATCAAGCAGACAGACATTAACATCTACTCCACAGCA 1443  
QY 1914 atcaactatgtccgtctccttgtatltcaagaagatltaccaaaagttgtaaaactltgaa 1973  
|||||  
Db 1444 ATCACTGATGTCCCTCTCTTGTATTAAGGAGATTACCAAAAGTGTAATAACATTTGAG 1503  
QY 1974 gatttcccaattctgtccaggaagaatltcaaatataaagtgaagcagtgagtaagaat 2033  
|||||  
Db 1504 GATTTTCCAATTCGCCAGAGGAATATTCAAAATTAATGAAAGTACGTGTAAGAT 1563  
QY 2034 gggcnaactaaatcaagatcctcgtgtccctgacccgctactactactagtttcgttaatat 2093  
|||||  
Db 1564 GGGCCAACATAATCAGATACCTCGGTGCTGACCCGCTATTCCTAGTTGTTGTAATATG 1623  
QY 2094 gaagagaatcaggttcaagacatattgtggccctctcctcaatctgctcaaaagaatctgta 2153  
|||||  
Db 1624 GAGAGAGATCTAGGCTTGAGACTCATTTGGCCCTCTCCATCTGCTACAAAGAAATCTGTA 1683  
QY 2154 gatcaagaagaagaacagataatgttcaagaagaaggaatgtcatccgttctctgtatlt 2213  
|||||  
Db 1684 GATCAAAAGAGAAACAGATATATGTACAGACAAAGGAATGTCAATCTGTTTCTGTATTT 1743  
QY 2214 gatgagaacccaagctgtgtacatcacaagaagaatatatacaagcttcttccccaatccagct 2273  
|||||  
Db 1744 GATGAGAACCGAAGCTGTGTACTCTCAGAGAAATATACAGCTTCTGCCCAATCCAGCT 1803  
QY 2274 ggaagtcagcttgaagatccaaggttccaaagctcccaatcatatgacaagaatcagtcagc 2333  
|||||  
Db 1804 GGAAGTCAAGCTTGAGGATTCAGAGATTCOAAGCTTCAACATCATGACAGCATCAATAGGC 1863  
QY 2334 tatgttttataatgtgaggtgtcaggttgttggatagaggtgtgatactgttaact 2393  
|||||  
Db 1864 TATGTTTGTATGATGTTGCAAGTTGTGATGTTGTTGATGAGGTGAGATGATGATGATGAT 1923  
QY 2394 ctaagcatltggaacagactgaacttctctgtcttctctctctctgtgatalacttcaaa 2453  
|||||  
Db 1924 CTAGCATTTGGAGACAGAGACTGACTCTTCTGCTCTCTCTCTCTGATATACCTTCAAA 1983  
QY 2454 cacaaaatgtctataatgaagacaactcaacccatctccatctcaggaagaacatgtcttc 2513  
|||||  
Db 1984 CACAAAATGTGCTATGTAAGACACACTCACCTTATTTCCATTTCTCAGAGAAAGTGTCTTC 2043  
QY 2514 atgtcagtgtaaaaacccaggttctatgtatctctgtggtgcacaacactcagacttgcgaac 2573

```

Db 2044 ATGTCGATGGAAAAACCCAGGCTCTATGATTCTGGGGTGGCCACACCTCAACACTTTTGGGAAC 2103
Qy 2574 agagcagtcacgcgcctctctgaaggttctctagttgtgacaagaacacgtgtgataltaac 2633
Db 2104 AGAGGCAATGACCGCTTACTGAGGTTTCTAGTTGTGACACACACACTGCTGATTTATTTAC 2163
Qy 2634 gaggaacagltla tgaagatattcaagacatactgtctgagtaaaaaaactgcatltgaacca 2693
Db 2164 GAGGACAGTTATGAAAGATATTTCACGATACCTGCTGAGTAAAAAACATGCCATTGAAACA 2223
Qy 2694 agaaagcttccccaagaattcaagacacccctagcactaggcaaaagcaatttaattgcacc 2753
Db 2224 AGA----- 2226
Qy 2754 ccaccagctctgaaagccatcaacgcgggaataactcgtactactctctcagtcagataca 2813
Db 2227 -----GAAATACTCGTACTACTCTTTCAGTCAGATCA 2259
Qy 2814 gaggaatltgactatgtatataccataaccagltgaaatgaaagaagatlttgacatc 2873
Db 2260 GAGGAAATTTGACTATGATGATACCATATCAGTTGAAATGAAGAAGAGATTTTGACATT 2319
Qy 2874 tatgataagagttgaaataaccagagcccccgcgcagcttcaaaaagaaaaacagacacatttc 2933
Db 2320 TATGATGAGGATGAAATCAAGCCCGCACCTTTCAAAAAAGAACAGACACTATTTT 2379
Qy 2934 atctctgagtgagagagcctcgtgagatattgagatgagtagctcccaacagltcttaaga 2993
Db 2380 ATTGCTGAGAGTGAGAGGCTCTGGGATTTATGAGATGAGTAGCTCCCGCACATGTTCTAAGA 2439
Qy 2994 aacagagctcagagtgagcagtgcccccaagltcaagaagatgtgttccagaatttact 3053
Db 2440 AACAGGCTTCAGAGTGAGGAGTCCCTCAGTTCAGAAAGTGTGTTTCCAGAAATTTACT 2499
Qy 3054 gatgcctccttactacgccccttaaccgttgagaactaaatgaaacttttgagactccg 3113
Db 2500 GATGGCTCTTACTACACCCCTTATACCGTGAGAACTPAATGAACATTTGGAGACTCCG 2559
Qy 3114 gggcatalataagaagcagaagltgaaagataacatcaglttaacttcagaatcagcc 3173
Db 2560 GGGGCATATATAGAGCAGGAAGTTGAAGATATATCATGTAACTTTTCAGAATCAGGCC 2619
Qy 3174 tctgtccctatcccttactatctagccttacttcaatgagaagatcagaagga 3233
Db 2620 TCTGTGCTCTTCTCTCTTATTTAGCCTTATTTCTATGAGAGATCAGAGCAAGGA 2679
Qy 3234 gcgaagactgaaaaaacttgtaagccttaatgaaccaaacttactttggaaagtg 3293
Db 2680 GCAGAACCTTAGAAAAAATTGTCAAGCTTAATGAACCAAAAACCTTCTTTGGAAAGTG 2739
Qy 3294 caacatcatatgacacccactaaagatgagtttgaactgcaaaagccttggtctattctct 3353
Db 2740 CAACATCATATAGGCAACCCACTAAAGATGAGTTTGACTGCAAAAGCCTGGCTTATTTCTCT 2799
Qy 3354 gatgttgaactggaanaagatgtgcaactcagagcctgatttgacccctctgtgtcgcac 3413
Db 2800 GATGTTTACCTGGAAAAAGATGTGCACCTCAGGCCGTGATGGACCCCTCTGCTGTCAC 2859
Qy 3414 actaagcagtaaacccctgcctcatatgagagacaagtgacagatagaagaatttgctctgttt 3473
Db 2860 ACTAACACACTGAACCCGCTCATGTGGAGACAAAGTACAGATCAAGAAATTTCTCTGTTT 2919
Qy 3474 t---tcaacatcttltgatagaacaaagcctggtactcactgaaataatgaaagaac 3530
Db 2920 TTCTCTACACATCTTTGATGAGAGACCAAAAGCTGCTTCTCAGTGAATAATGAAAGAAAC 2979
Qy 3531 tgcagagctccctgtcaataatccagatggaagatcccaactttaaagagaatatactgcctc 3590
Db 2980 TGCAGGGCTCCCTCAATATATCAGATGGAAGATCCCACTTTTAAAGAGAAATATATGCTTTC 3039
Qy 3591 catgcaatcaatgtgctaatatagtatactactggtcttagtaatgtcttaagatcaa 3650

```

---

```

Db 3040 CATGCAATCATGCTACTAATAATGATACACTACCTGCTTATGATAGGCTCAGACATCA 3099
Qy 3651 aggattcga tgg latctgtctcagcatgggcagcaatgaaacatccattcatcatttc 3710
Db 3100 AGGATTGATGGTATCTGCTCAGCATGGCAGACATGAAACATCATTTCTATTTATTTTC 3159
Qy 3711 agtggacatgtgtcaactgtacgaaaaaagagagatlaaaatggtcagctgtacaatctc 3770
Db 3160 AGTGAACATGTGTTCACTGATACGAAAAAAGAGAGGTATTAAGATGGCACTGTACAATCTC 3219
Qy 3771 tatccagtggttttttgagacagtggaatgttataccatccaagctggaatttgccggtg 3830
Db 3220 TATCCAGGTGTTTTTGAACACTGGAAATGTTTACCATCAAGCTGGATTTTGGCGGGTG 3279
Qy 3831 gaatgcctatttggcagcatcatalatgctggatgagacaacttttctgtgtgacagc 3890
Db 3280 GAATGCTTATTTGGGAGAGATCTTACATGCTGGATGAGACACACTTTTCTGTGTACAGC 3339
Qy 3891 aataagtgtaagactcccttggaaatggtctcttgacacattagaatlttcagattaca 3950
Db 3340 AATAAGTGTCAAGACACCCCTGGGAATGGCTCTGACACATTAAGATTTTCAGATTACA 3399
Qy 3951 gcttgaagaacaaatggaacagtgggcccaaaagctggccgaacttcatctatccggatca 4010
Db 3400 GCTTCAGACAAATATGAGACAGTGGCCCCAAAGCTGGCCAAAGCTTATTTCCGGATCA 3459
Qy 4011 atcaatgtccttgagacacaaagagcccttctcttgatcaaaagtgagatctgttgacca 4070
Db 3460 ATCAATGCTGTGAGCACACAGAGAGCCCTTTCTTGATCAAGTGAGATCTTGTGGACCA 3519
Qy 4071 atgattatltcacgcatcaagaacccaaggtgtccgctgaagaaglttccagccttcaatc 4130
Db 3520 ATGATTATTTACAGGATTAAGAACCCAGAGTGCCCTCGAAGTGTCTCCAGCCTCTTACATC 3579
Qy 4131 tctcagttatacatatlatagttcttgatggaaagagtgagagactatccaggggaat 4190
Db 3580 TCTCAGTTTATATATGATGATATGATCTTGTGATGGAGAAAGTGGACACTTATTCAGGAAT 3639
Qy 4191 tccacttgaaccttaatgtgtctctcttcttgccaatgtggaatcaatctggaataaaacaaat 4250
Db 3640 TCCACTGGAACTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3699
Qy 4251 attttaaaccctccaaattatgtctgtatataatccgtttgcaaccaatcatataagatc 4310
Db 3700 ATTTTAAACCTTCACATTAATGCTGATACATCCCTTTGCAACCAACATTAATGAGATT 3759
Qy 4311 gcgaagcacttcgcgatggaatgtgaatggcggtgtgattaaatgttgaagatccatgt 4370
Db 3760 CCGACACTCTTGCATGAGATGAGTGTGATGGGCTGTGATTTAATAGTTGACGATGCCATTG 3819
Qy 4371 ggaatggaagtaagaacaaatatacaagatgcacagatctactgtctcaatccatcttacaat 4430
Db 3820 GGAATGGAAGATTAAGCAATATACATGACATGACAGATTACTGCTCATCTTATTTACCAAT 3879
Qy 4431 atgtttgcaactgtgtctcccttcaaaagctggaactcaactcccaaggaggaatgaatgcc 4490
Db 3880 ATGTTTGGCACCTGTGCTCTTCAAAAGCTGACCTTCACTCCAAAGGAGAGATTAATGCC 3939
Qy 4491 tggaaacctcgaatgaaataatcaacaaagatgtgtgcaagtggaagcttccagaagaatgt 4550
Db 3940 TGGAGACCTCAAGTGAATATATCAAAAGAGTGCGTGCAGAGTGACCTTCAGAAAGCAATG 3999
Qy 4551 aaagtacagagtaactactcaaggagtaaaatctctgtcttaccagcatgtatgtgaag 4610
Db 4000 AAAGTCACAGAGTAAATACATCAACGAGTAAATATCTGCTTACACAGATATATGTGAAG 4059
Qy 4611 gagttccatctccagcagatcaagaatggccaatcaagtggaactctcttttccagaatggc 4670
Db 4060 GAGTTCCTCATCTCCAGAGTCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 4119
Qy 4671 aaagtaaaagtttccaggaataaagaactccttcaaacctgtgtgtgaactctctagac 4730
Db 4120 AAAGTAAAGGTTTTTCAAGGAATCAAGACTCTTCAACACTGTGTGTAACCTCTTAGAC 4179

```

QY 4731 ccaccgttaactcgtccttaactcgaatccacccccagaattcgggtgcacagattgccc 4790  
|||||  
Db 4180 CCACCGTTACTGCTGCTACTCTTCGATTTACACCCACAGATTGGGTGACACGATTTGCC 4239

QY 4791 ctgaagatggaggttcctgggtcgcagagcaagacctctac 4832  
|||||  
Db 4240 CTGAGATGGAGGTCTGGGCTCGAGGACACAGACCTCTAC 4281

RESULT 5  
LOCUS 108642 108642 4548 bp PAT 02-DEC-1994  
DEFINITION Sequence 2 from Patent WO 8800831.  
ACCESSION 108642  
VERSION 108642.1 GI:588648  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4548)  
AUTHORS Pasek,M.P.  
TITLE DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE POLYPEPTIDES IN HIGH YIELDS  
JOURNAL Patent: WO 8800831-A 2 11-FEB-1988;  
FEATURES  
source 1..4548 Location/Qualifiers  
BASE COUNT 1344 a 1004 c 1002 g 1197 t 1 others  
ORIGIN

Query Match 51.4%; Score 4085; DB 10; Length 4548;  
Best Local Similarity 95.1%; Pred. No. 0;  
Matches 4325; Conservative 0; Mismatches 31; Indels 192; Gaps 3;

QY 477 gccaccagaagatactactcctggtcagtcgaggaactgtcatggagcatatgcaagtgat 536  
1 GCCACCAAGAAAGTACTACTCGGTGCACTGAGAACTGTCAATGATATGCAAAAGTGAT 60  
Db 537 ctctggtgagctgctgttgagcgaagaattcctcctcctagagtgccaaatcttctcattc 596  
|||||  
Db 61 CTGGTGAAGTGGCTGTGGAGCGCAAGATTTCTCTAGAGTGCCAAATCTTTTCATTTC 120

QY 597 aacaactcagtcgtgtacaacaaagactcgtttgtagaattcaagggttcaacctttcaac 656  
|||||  
Db 121 AACACCTCAGTCTGTACAAAAGAAAGACTCTGTTGTAGAAATTCAGGATCACCTTTTCAAC 180

QY 657 atcgctcaagcgaagggccaccttgatggtgtcgttaggtcctcctacacatccagagcttgaagt 716  
181 ATGCGTAAGCCAAAGGCCACCTGATGGGTCTAGGTCCTAGCCATCCAGGCTGAGGTT 240  
Db 717 tatgatcagttggtlcaatcaacttaagaacatggtctccatcctgtlcaagttctcattgct 776  
241 TATGATACAGTGGTCACTTAACATTAGAACATGCTTCCATCCTGTGCAAGTCTTCAATGCT 300

QY 777 gttgtgtatcctcactcgtgaagactctctgagggagctgtaatagtatcaagcagtcgaa 836  
|||||  
Db 301 GTTGGTGTATCTCACTACGAAAGCTTCTGAGGGAGCTGAATATGATATCATGACACATCA 360

QY 837 agggagaagaagatgataaagctctccctggtgtggaagccatacatatgtctggaggtc 896  
|||||  
Db 361 AGGGAAAGAAAGATATATAAGTCTTCCCTCGTGGAGAACCATATATGTCTTGGCAGGTC 420

QY 897 ctgaagaagaatggtccaatggtccctcgtgacccactgtgcctlaactactaatactcttct 956  
|||||  
Db 421 CTGAAGAAGATGCTCAATGAGCTCTGAGCCCACTGTGCTTACTACTACATATCTTTCT 480

QY 957 catgtgacctgtgttaaaagacttgaattcaagcctcattggagcctcactagtatgata 1016  
|||||  
Db 481 CAGTGGACCTGGTAAACACTTTGAATTCAGGCTCATTTGGACCTCTACTAGTATATAGA 540

QY 1017 gaaggagtcctgcgaaggaagaagacacagaccttgcacaaatttaactacttttgc 1076  
|||||  
Db 541 GAAGGAGTCTGGCCAAAGAAAGACACACACTTGCACAAATTTATACACTTTTTCGT 600

QY 1077 gtaattgataaaggaaagtctggcaactcgaacaaagaactcctctgaltgcagatag 1136  
601 GTATTGATGAAGGAAAGTGTGGCACTCAGAAACAAAGAACTCTTGATGACGAGATAG 660

QY 1137 gatctgcatactgtctggggtcgtgacctaaatgcaacagtcgaatggttatgtaaacagg 1196  
661 GATGCTGCATCTCTCGGGCTGGCCCTAAATGTCACACACTCAATGATGATTAACAGG 720

QY 1197 tctctg---ccaggtctgattgtagtcacagaagaatcagtcattcattggcagtgattgga 1253  
|||||  
Db 721 TCTCTGCTACCAAGTCTGATTGGATGCCACAGAAATCACTTAATTTGGCATTTGGA 780

QY 1254 atggcaccacactcctgaaagtgcactcaatattcctcgaagtgacacatctctgtgag 1313  
781 ATGGCACCACTCCTGAAAGTGCACCTCAATATTCCTCGAAGGTGCACACATTTCTTGAGG 840

QY 1314 aaccatcgccaggtctcctgtggaatctgcgcaataacttctcctactctcaaacatc 1373  
841 AACCATCGCAGCGCTCTTGAATCTCGCAATACTTCTCTACTGCTCAAAACCTC 900

QY 1374 ttgatgaccttggagaagttctactgtttgtgcatactctcccaacaatgattggtc 1433  
901 TTGATGACCTTTGGACAGTTTCTACTGTTTGTGCTATATCTCTTCCACCAACATGATGTC 960

QY 1434 atggaaactatgtlcaaaagtagacagctgtccagaggaaccccaactagcaatgaaat 1493  
961 ATGGAAGCTTATGTCAAAAGTAGACAGCTGCCAGAGAACCCCACTCAAGATGAATAAT 1020

QY 1494 aatgaagaagcgaagaagactctgtagatgatacttactgattctgaatagtgtgtcagg 1553  
1021 AATGAAGAAACGGGAAGACTATGATGATGATCTTACTGATTTCTGAATGATGGTCAAG 1080

QY 1554 ttgatgatcaaacctctcctcctatccaaattgcgtcagttgcgaagaagatcct 1613  
1081 TTGATGATGATCAACCTCTCTCTTTATCCAAATTCGCTCATTTGGCAAGAAAGATCCT 1140

QY 1614 aaaacttgggtacattacattgtcgtcgaagaggaagacttggagactgtcccttagtc 1673  
1141 AAAACTTGGGTACTATTAATTTGCTGGAAGAGAGAGACTGGGACATATGCTCTTAGTC 1200

QY 1674 ctgcgcccgatgacgaagaattataaagtcaatattgacaatggtccctcagcgatt 1733  
1201 CTGCCCCCGATGACAGAAATTTAAAGTCAATATTTGAACAATATGCCCCGCGGATT 1260

QY 1734 ggtaggaagtacaaaaaagtcggaattatggaacacagaatgaacacttgaagactcgt 1793  
1261 GGTAGGAAGTACAAAAAGTCCGATTTATGCGATTAACAGATGAACCTTTAAGACTCGT 1320

QY 1794 gaagcattcgaatgaaatcgaatccttggagacttacttatacttggagagttggagac 1853  
1321 GAAGCTATTTCAGCATGAATCAGGAATCTTGGACCTTACTTTATGAGGAGTTGGAGAC 1380

QY 1854 acacgttgtatataatlaagaatcaagaagcagagaccataataactctacccctcagga 1913  
1381 ACACGTGTGATTTATTTAAGTAATCAAGCAAGCAACATATTAACATCTTACCTCCGGA 1440

QY 1914 atcacgtatgtccgtccttgtatctcaagagagatcccaaaaggtgtlaaacaattgaa 1973  
1441 ATCACGTGATGTCGCTCTTTGTATTCAGGAGAGATTACCAAAAGGTGTAACATTTGAAG 1500

QY 1974 gatttccaattctgcagagagaataatcaatataaalygacagtgatgataagat 2033  
1501 GATTTTCAATTTCTGCAGAGAAATATTCAAATATAAATGACACTGATGATGAAGAT 1560

QY 2034 gggccaactaaatacaatcctcgtgtgcgcagccgctatctactctgtatgtaata 2093  
1561 GGGCCAACATTAATCAATCCTCGTGTGACCCGCTATTTACTGTAGTTTCTGTTAATATG 1620

QY 2094 gagagagatcagcttcaagacatcttggtccctctcctcactcgtcaaaagaatctgta 2153



QY 4125 tacatctcagttatcatcatatagttcttgaggaaagtggcagactatga 4184  
|||||  
Db 3841 TACATCTCTCATGTTATCATATATATATAGTCTTATGGGAAAGTGGACACTTATCGA 3900  
QY 4185 ggaattccacttgaaccttaatgtctctcttttgcaatgtgatactcctgggataaa 4244  
|||||  
Db 3901 GGAATTCACATGGAACCTTATATGCTCTTTTGGCAATGTGGAATTCATCTGGGATATAA 3960  
QY 4245 cacaatatttcaacctccaattatgtctgatacacaacgtttgcaccaactaatat 4304  
|||||  
Db 3961 CACATATTTTAAACCTCCATATTTATGCTGATACATCCCTTTGACCCAACTCATATAT 4020  
QY 4305 agcatctcgaacctcttctgcaatgagttgagtggtgctgagtttaaatagttgagatg 4364  
|||||  
Db 4021 AGCATGTGCACACTCTCTTCGATGAGTGTGATGGGCTGTGTATTAATAGTTGACAGCTG 4080  
QY 4365 ccattgggaatgagagtaaaagcaataatcagatgcagagatctactgtctactactt 4424  
|||||  
Db 4081 CCATTGGGAATGAGAGTAAGCAATATACATGACAGATTACTGCTTCACTCTACTTT 4140  
QY 4425 accaatatgttgcacctgtgtctcttcaaaagctgcacttcaacctccaaggagagt 4484  
|||||  
Db 4141 ACCAATATGTTTGGCCACTGTCTCTTCAAAAGCTGACTTCACTTCAAGGAGGAGAGT 4200  
QY 4485 aatgcctgggaacctctcagtgtaataatccaaaagtgctgagtgagtgagttcagag 4544  
|||||  
Db 4201 AATCCCTGGGAACCTCTAGGTAAATTAATCAAAAGAGTGGCTGCAAGTGGAGATTCCAGAG 4260  
QY 4545 acaatgaagatcacaggaagtaactactcaagagtaaaatctctgtctaccagatgtat 4604  
|||||  
Db 4261 ACATGAAAGTCAACAGAGTAACACTACAGGAGTAATAATCTGCTTACAGAGATGAT 4320  
QY 4605 gtgaagagttcctcatctccacagcagtaagaatgagtgagtgagttccttcttcaag 4664  
|||||  
Db 4321 GTGAGAGAGTCTCATCTCAGCAGTCAAGATGGCCATCAGTGGAGCTCTCTTTTTCAG 4380  
QY 4665 aatggcaaatgaaggtttttcaggaaatcaaacctctcacactgtgtgtaactct 4724  
|||||  
Db 4381 AATGGCAAGTAAGAGTGTTCAGAGGAATCAAGACTCTTCAACCTGTGTGTAACCTT 4440  
QY 4725 ctgaaccacacgttactgactcgtactcttgaatltcaaccacagagttgtgtcacag 4784  
|||||  
Db 4441 CTAGACCCACGTTACGTACGCTGCTACTCTTGATTCACCCCGAGAGTTGGTGCACAG 4500  
QY 4785 atgtccctgagagtgagagttctgtgctgcgagggacagagcctctac 4832  
|||||  
Db 4501 ATTGCCTGAGAGATGAGGTCTGTGGCTGGAGGACACAGGACCTCTTAC 4548  
RESULT 6  
LOCUS 108641 4551 bp PAT 02-DEC-1994  
DEFINITION Sequence 1 from Patent WO 8800831.  
ACCESSION 108641  
VERSION 108641.1 GI:588647  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4551)  
AUTHORS Pask, M. P.  
TITLE DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR  
VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE  
POLYPEPTIDES IN HIGH YIELDS  
Patent: WO 8800831-A 1 11-FEB-1988;  
JOURNAL location/Qualifiers  
FEATURES 1..4551  
BASE COUNT 1345 a 1003 c 1004 g 1198 t 1 others  
ORIGIN

Query Match 51.4%; Score 4085; DB 10; Length 4551;  
Best Local Similarity 95.1%; Pred. No. 0;  
Matches 4325; Conservative 0; Mismatches 31; Indels 192; Gaps 3;  
QY 477 gccacccgaagaataactactgtgtcagtggaactgtcatalggaactatagcaagatgat 536  
|||||  
Db 4 GCCACACGAAGATATCTACTGCTGGTGACGTGAGACTGTCATGTCATGATATGCAAGATAT 63  
QY 537 ctggtgagctgtcctgtgtgagcgaagatttccctcagagtgccaaaactcttccatctc 596  
|||||  
Db 64 CTCGGTGAAGCTGCTGTGGAGCAGAGATTCTCTCTAAGTGCAGAAATCTTTTCCATATTC 123  
QY 597 aacacctcagtcgtgtacaaaagacctgttcttgaatccagagttcaccttccaac 656  
|||||  
Db 124 AACACTCAGTGTGTACAAAAGACTTGTGTTAGATTCACAGGATCACCTTTTCAAC 183  
QY 657 atcgttaagccaaagccaccccttgatggtgtcgtgagttccatccatccacagctgaagt 716  
|||||  
Db 184 ATCGCTAAGCCAGGCCACCTGGATGGGTCTGCTAGGTCCTACATTCACAGGTGAGGTT 243  
QY 717 tatgatacagtggtcatttaacacttaagaacatggtctccatctctgtcagttcctcaagct 776  
|||||  
Db 244 TATGATACAGTGTGATTAACACTTAAGAACATGTGCTTCCATCCGTGCAGTCTTCATGCT 303  
QY 777 gtgtgtatctactctgaagaagcttcttgaaggagctgaatatgatatcagaacagttcaa 836  
|||||  
Db 304 GTTGCTGATCTCTACTGGAAGCTTCTGAGGAGCTGAATGTGATCAACACAGTCAA 363  
QY 837 agggagaagaagaatgaataagttctccctgttgaagccataatagtctcgaagctc 896  
|||||  
Db 364 AGGAGAAAGAAAGATGAATGAATAGTCTTCCCTGTGGAAGCCATACATATGCTGGAGGTC 423  
QY 897 ctgaagaagaatgtgtccaaatgagctctgaagcctgtgagttccttccatctatcttct 956  
|||||  
Db 424 CTGAAGAGAAATGTCCAAATGCTGACCTGACCCACTGTGCTTACTCTCATATCTTACT 483  
QY 957 catgtgacacctgtgaagaagcttgaattcaaggtcttcttgaagccactcagttatgaga 1016  
|||||  
Db 484 CATGTGACCTGTGAAGACTTAATTCAGGCTCATTTGAGCCCTACTAGTATGTAAGA 543  
QY 1017 gaaaggagctctgcccagaagaagaacacagccttgcacaaattataactcttctgt 1076  
|||||  
Db 544 GAAGGAGTGTGGCCAGGAAGAAACACACACTTGCACAAATTTATATCTCTTTTCT 603  
QY 1077 gtaattgataaggaagaaagtgtgcaactcagaacaaagaactcctgtatgagagatagg 1136  
|||||  
Db 604 GTATTTGATGAAGGAAAGATTTGGCAGCTCAGAAACAAAGAACTCTTGATGACAGATAGG 663  
QY 1137 gatgtgcatctgtctgtgagcctgtgcctaaatgagacagtcgaatgtgtatgtataacagg 1196  
|||||  
Db 664 GATCTGCATCTGCTCGGGCTTGGCTTAATATGACACAGTCAATGGTTATGTAAACAGG 723  
QY 1197 tctctg---ccaagttctgattgtgatagtccacagaagaatcagttcattgtgagatgtgaga 1253  
|||||  
Db 724 TCTCTGTACACAGGTGTGATGTGATGACACAGGAATACATCTTGTGGCATGTGATGGA 783  
QY 1254 atgggacccactcctgtgaagtgacactcaatatctctcgaagtgacacattctctgtgagg 1313  
|||||  
Db 784 ATGGGACACACTCTCTGAAGTGCACCTCAATATTTCTCGAAGGTCAACATTTCTGTGGAG 843  
QY 1314 aaccatgcgaagtgctcttgaagaatctgcgaataacttccctactgtgtccaaactc 1373  
|||||  
Db 844 AACCATGCGACGCGCTCTTGAATCTGCGCAATTAATCTTCTTACGCTCAACACATC 903  
QY 1374 ttgatacacttgagagtttctactgtttgttataatactcttccacccaacatgagtgagc 1433  
|||||  
Db 904 TTGATGAGCCTTGGACAGTTTCTACTGTTTGTGCAATATCTTCCACCAACATGATGGC 963  
QY 1434 atggaagcttatgtcagaatgacagctgtccagaggaacccaactcagaatgagaat 1493  
|||||  
Db 964 ATGGAAGCTTATGTCAAAAGTATGACAGCTGTCCAGAGGAACCCCAACTACGATGAATAAAT 1023  
QY 1494 aatgaaagaaggaagactatgatatgatatcttactgtatcttgaatgagatgtgtcagg 1553



```
|||||
Db 1024 AATGAAGAGGGAAGACTATGATGATCTACTGATTCGAATGATGTGTCAGG 1083
Oy 1554 ttgatgatagaactcccttccttccaaatcgctagcttgccagaagacacct 1613
Db 1084 TTTGATGATGACACTCTCCCTTCCTTTATCCAAATTCGCTCGATTGCCAAGAACGATCCT 1143
Oy 1614 aaacttggatcatatacttgcctgaagaaggagactggagactatgctccctagtc 1673
Db 1144 AAAACTGGGATACATTAATTGCTGCTGAGAGGAGACTGGAGCTATGCTCCCTTAGTC 1203
Oy 1674 ctgcgcccgatgacagaagtataaagtcaatatltgaacaatggccctcaaggatc 1733
Db 1204 CTCGCCCCCGATGACAGAGTTATAAAGTCAAATTTGGAACAAGGCCCTCAGGGATT 1263
Oy 1734 ggttgggaagtacaataaagtcgatlttatgcaatacagaatgaaaccttaagactgt 1793
Db 1264 GGTGAGAGATGCAAAAAAGTCGATTTATGGCATACACAGATGAACCTTTAAGACTGCT 1323
Oy 1794 gaagcatatcagcatgaatcaggaatccttgggaccttacttatacttggggaagtggagac 1853
Db 1324 GAACCTATTCAGCATGATGATCAGGAATCTTGGGACCTTACTTTATGGGGAAGTTGGAGAC 1383
Oy 1854 acaactgttgaattataatgaagaatcaagcaagcagacataatacattaccctcaagga 1913
Db 1384 ACACGTGTGATTATATTTAAAGATCAACCAAGCACACATATACATCTACCCCTCACGGA 1443
Oy 1914 atccatatagtccgtcccttctgtatcatcaaggagatfaccaaaagtgtaaaactttgaag 1973
Db 1444 ATCAGCTGATTCGCTCCCTTTTGTATTCAAGGAGATTACCAAAAAGGTGTAACATTTGAAG 1503
Oy 1974 gatttcccaatcttgcagagagaatataccaataataatggacagctgactgtaagaagt 2033
Db 1504 GATTTTCCAATTTGCGCCAGAGAAATATTCAAAATATAATGACAGTGCTGTAGAGAT 1563
Oy 2034 gggcccaactaaatcaagatccctcggtgccttgacccgctattactagtttgtaatatg 2093
Db 1564 GGGGCAACTAATACAGATCCTCGGTGCTGACCCGCTATTACTCTAGTTTCCTTAATAG 1623
Oy 2094 gaggagatcatagcttcaaggactcatggccctcccaatctgtctatacaagaactgtta 2153
Db 1624 GAGGAGATCTTAGCTTCAGGACTCATTTGGCCCTTCCTCATCTGCTACAAAGAACTGTGA 1683
Oy 2154 gatcaaaaggaaacagataatgtcagacaagaaggaaatgtcaatccctgttcttctglatc 2213
Db 1684 GATCAAAAGAGAAACACATATATGTCACAAAGAGAAATGTCATCCTGTTTCTGTATATT 1743
Oy 2214 gatagaagacgaagcttggtaacctcaagaagaataatacaagcttcttcccacaatccagct 2273
Db 1744 GATGAGAACCCGAACCTGCTACCTCACAGAAATATTAACAGCTTTCTCCCAATCCAGCT 1803
Oy 2274 ggaatgcaagcttggagatccagaagtltccaagcctccaacatcatgacagcatcaatgac 2333
Db 1804 GGAATGCAGCTTGAGGATCCAGAGTTCCAAAGCCTCCAAATCATATGCACAGCATCAATGGC 1863
Oy 2334 tatgttlttgaatgttgcagltgtcagltgttgcataaggtgggcatctactgtaact 2393
Db 1864 TATGTTTTGATGATTTGCAATTTGCAATTTGTTGTCATGAGTGGCATCTGATACATT 1923
Oy 2394 ctaagacttggagacagactgacttcttctgtcttcttcttctcggatataccctcaaa 2453
Db 1924 CTAAGCATTTGGAGCACAGACTGACTTCCTTCTGTCTTCTCTCGGATATACCTTCAAA 1983
Oy 2454 cacaanaatggtctatgaagacacactcccatatcccatltcagaaggaaactgtcttc 2513
Db 1984 CACAAAATGGCTTATGAAGACACACTCACCTATTTCCATTTCTCGAGGAACCTGTCTTC 2043
Oy 2514 atgtcgaatgnaaaacccaggtctatggaatctctggggtgccaaactcagacttgcgaac 2573
Db 2044 ATGTCGATGGAACCCAGGCTATGATGATCTGGGGTGCCAACTCAGACTTTCGGAAC 2103
Oy 2574 agaggaatgaagcccttaactgaaaggttctcagttgtgacaagaacaactcgtgataactaac 2633
|||||

|||||
Db 2104 AGAGCAGTACCGCCTTACTGAAGGTTTCTAGTTGTGACAAAGAACACTGCTGATTTATTAC 2163
Oy 2634 gaggacagttatgaagataatcacaactactgtgcagtgaaataaacaatgccattgaaca 2693
Db 2164 GAGGACAGTTATGAAGATATTTTACACATATCTTGCTGATTAATAAACAAATGCCATTGAAACA 2223
Oy 2694 agaagcttctccagaat----- 2711
Db 2224 AGAAGCTTCTCCAGAGATCCTCTGCTTGATTAACCACTATGTAAGTCAAGATACAAAA 2283
Oy 2712 ----- 2711
Db 2284 GAAGAGTGAATCCCAAGAGAGTCCACCAAAAAACAGCTTTTAAGAAAAAGGATACC 2343
Oy 2712 ----- 2711
Db 2344 ATTTGTTCCTTAAGCTTTTGAAAGCAATCATGCAATAGCAGCAATTAAGAGGACAA 2403
Oy 2712 -----tcaagacaacctagcactaggcaaaagcaatltaat 2747
Db 2404 AATTAAGCCCGAATATGAAGTCACTGGGCAAAAGCAAGGTAGACTGAAGAGCTGTGCTCT 2463
Oy 2748 gccaccccaaccagttcttgaagccatacaogggaaataaactcgttactactcttcagta 2807
Db 2464 CAAAACCCACAGCTTTGAACGCCATCAAGCGGAATTAACCTGTAATCTTCAGTCA 2523
Oy 2808 gatcaagaagaaatgactatgatataccatatcagtttgtaaaatgaagaagaatttt 2867
Db 2524 GATCAAGAGAAATTTGACTATGATGATACATACATACATACGTTGTAATTAACAAAGAAATTTT 2583
Oy 2868 gacattatgatgagatgaagaaatcaagagcccccagcttccaagaagaacaagacac 2927
Db 2584 GACATTTATGATGAGAGATGAAGAAATCAGAGCCCGCACAGCTTTCAAAAGAAACAGCAC 2643
Oy 2928 tatlttatgtgcagagtgagagagctctggattatggagatgaagtagtcccccaatgltc 2987
Db 2644 TATTTTATTTGCTGAGTGAGAGGCTGTGGATTTATGGATGAGTAGTCCCCACATGTT 2703
Oy 2988 ctaagaaacagggctcaagatggcagtgccctcaagttcaagaagaagtgttctccagaa 3047
Db 2704 CTAAGAAACAGGGCTCAGAGTGGCAGTCCCTCAGTTCAAGAAAGTTTTCACAGGA 2763
Oy 3048 ttactatgagctccttactcaagccctataacogtggagaactaaatbgaacatttggga 3107
Db 2764 TTTACTGATGGCTCCTTACTACAGCCCTTATACGCTGAGAAATTAATGAACATTTGGGA 2823
Oy 3108 ctccctggggccaatataaagagcagaagtltgaagaataatcatatgtaacttcaagaat 3167
Db 2824 CTTCTGGGGCCATATATTAAGACGACAAGTTGAAGATTAATCATGATGTAACCTTCAGAAAT 2883
Oy 3168 caggcctcgtccctatcttcccttactatctagccttatttcttatatgaggaatcaagag 3227
Db 2884 CAGGCTCTCGTCCCTATATTCCTTATTTCTAGCCCTTATTTCTTATNGAAGAAATCAGAGG 2943
Oy 3228 caaggagcaaacctagaaaaaacttltcaagcctaaatgaacccaaacttactcttgg 3287
Db 2944 CAAGGAGCAGAACTTGAAGAAAACTTTGTCAAGCTTAATGAACCAAAACTTACTTTGGG 3003
Oy 3288 aaagtcaaatcatatagcccccacaaagaatgaatgttgccttcaagccttggactat 3347
Db 3004 AAAGTGAACATCATATGGAACCCCACTTAAGATGAGTTGACTGCAAAAGCCTTGAGCTTAT 3063
Oy 3348 ttctctgaatgttgaacttgaanaaagatgtgaactcaagcctgacttgaacccctctgttc 3407
Db 3064 TTTCTGTGATGTTGACCTTGGAAGAAAGATGTGACATCAGGCTGATTTGGACCCCTTCTGTGC 3123
Oy 3408 tgcacactaaacacatgaacccctgtcctcaatggagacaagtgtgacgttacaagaaattgtct 3467
Db 3124 TGCCACACTTAACACACTGAACCTGCTCATGAGGAGACAAGTGAAGATGAAGATTTGCT 3183
Oy 3468 ctgtttt---tcacacatcttgaatgagccaaagcctgttacttccacttgaanaaatatgaa 3524
Db 3184 CTTGTTTTCTTCACCATCTTTGATGAGACCAAAAGCTGTGTACTTCACTGAAAATATGGA 3243
|||||
```



|    |      |   |      |
|----|------|---|------|
| QY | 3525 | agaacatgcaaggtccctcgataatccagatggaagatcccaatttaaagaattat    | 3584 |
| Db | 3244 | AGAAATCTGACAGGCTCCCTCCAAATATCCAAATGGAAGATCCACTTTTAAAGAAATAT | 3303 |
| QY | 3585 | gccttcacatgaacatggtctcaataatgataacacacccggttgtaaigtcaag     | 3644 |
| Db | 3304 | CGCTTCCATGCANCAATGGCTACATATATGATGATACCTACCTGGCTATGATATGATG  | 3363 |
| QY | 3645 | gatacaagatctgattgtaactgctcaagatgtagcaagaatgaaacatccactta    | 3704 |
| Db | 3364 | GATCAAGGATTCGATGATATCTGCTCAGCATGGCAGCATGGAATGCAATCCATCTAT   | 3423 |
| QY | 3705 | cattcagtggaacatggttcacatgtaaaaaaagaagatataaaatgacctgac      | 3764 |
| Db | 3424 | CATTTCACTGGACATATGTTTCACTGATACGAAAAAAGAGAGATATATAATGTGCACT  | 3483 |
| QY | 3765 | aatcctcaacagtggttttgagaacagtgaaatgacatccaaagcttgaaattgg     | 3824 |
| Db | 3484 | AATCTCTATCCAGGTGTTTGTGAGACAGTGGAAATTTTACCATCCAAAGCTGGAA     | 3543 |
| QY | 3825 | cgagtggaatgacctatcttgacagacatctcaatgctggagatagacaactttctg   | 3884 |
| Db | 3544 | CGGGGGAATGCTTATTTGGCGAGCATCTATCATAGTGGATGAGCACACTTTCTG      | 3603 |
| QY | 3885 | tacagcaataagtgtagaactccctcgggaagtgctctcgacacataagaaattcag   | 3944 |
| Db | 3604 | TACAGCATATAGTGTGCACTCCCTCGGGAAATGGCTTTCGACACATTAAGATTTTC    | 3663 |
| QY | 3945 | attcagcttcaagacaataatggaacagtgtagcccaaaagctggcagaactcatatcc | 4004 |
| Db | 3664 | ATTACAGCTTCAGGACAAATTTGACACAGTGGGCCCAAACTGGCAGACTTCATTAT    | 3723 |
| QY | 4005 | ggatcaatcaatgctcggagacacaaaggaagcccttcttgataaaggtagactgtg   | 4064 |
| Db | 3724 | GGATCAATCAATGCTCTGGACACCAAGAGAGCCCTTTCTTGATCAAGTGATCTG      | 3783 |
| QY | 4065 | ggaccaatgattatctcaagcatcaagaagccagagtgcccgacagaattctcaagctc | 4124 |
| Db | 3784 | GCACCAATGATTATCTCCGGCATCAAGACCCAGGGTGGCCGTCAABAAGTCTCCAG    | 3843 |
| QY | 4125 | tacatctcagttatcaacatgtaagctgtagtggaagaagtggcagactatcga      | 4184 |
| Db | 3844 | TACATCTCTCAGTTTATCATCATGTAATAGCTTTGATGGGAATAATGGCAGACTTAT   | 3903 |
| QY | 4185 | ggaatctcactggaacctaataatggtctctcttgcaatgtagatcaactggaataaa  | 4244 |
| Db | 3904 | GGAAATTCACACTGGAACTTAATGATGCTCTTTGGCAATGTGATTAATCTGGGA      | 3963 |
| QY | 4245 | cacaaatattttaaaccctccaattatgctcgatataccglttgacccaactcatat   | 4304 |
| Db | 3964 | CACATATATTTTAACTCCCAATTAATTTGGTATCGATATCCGTTTACCCCAACTAT    | 4023 |
| QY | 4305 | agaaatcgagaactcttcgacatgagatgtagagggctgtagttaaataagttagcaag | 4364 |
| Db | 4024 | AGCAATTCGACACACTCTTCCCATGGAGTTGATGGCTGTGATTTAAATATGTTCCAG   | 4083 |
| QY | 4365 | ccattggaatgagagtaagaacatatacagatgacagatbactgcttcaactt       | 4424 |
| Db | 4084 | CCATTTGGGAATGGAGATTAAGCAATATCAGATGACACAGATTACTGCTTACCT      | 4143 |
| QY | 4425 | accaaatgtttgcaacactggtctccttcaaaaagctcgacttcaactccaaggagagt | 4484 |
| Db | 4144 | ACCAATATGTTTGGCACCTGTGCTCTTCAAAAGCTCGACTTCACCTCCAAAGGAG     | 4203 |
| QY | 4485 | aatgcctggagacctcagatgtaataatccaaaagatgtgctgcaagtgaactccaag  | 4544 |
| Db | 4204 | AATGCTGGAGACCTCAGTGTATTAATCCAAAAAGATGGCTCTCAAGTCCAGAG       | 4263 |
| QY | 4545 | acaatgaagtcacagagtaactactcagggagtaaaatctgcttaccagaatgat     | 4604 |
| Db | 4264 | ACAATGAAGATCACAGGATTAATCTACGAGGAGTAAATCTCTGCTTACCAGATGAT    | 4323 |

|  |  |  |             |
|--|--|--|-------------|
| QY   | 4605                                   | ggtgaagagatcttccatctccagcgatcaagatgagccatcagttgactctctttttcaag | 4664        |
| Db   | 4324                                   | GTGAGGAGATTCCTCATCTCCAGCACTCAAGATGGCCATCACTGAGCTCTCTTTTTCAG    | 4383        |
| QY   | 4665                                   | aatggcaaatgaagtttttctcaaggaaatcaagactccctacacctgtgtgaactct     | 4724        |
| Db   | 4384                                   | AATGGCAAAATGAAGTTTTCAGGAAATCAAGACTCCTTCACACTGTGTGAATCTT        | 4443        |
| QY   | 4725                                   | ctagaaccacgltactgaactcgtactccttcgaatcaccccaagatltggtygcacag    | 4784        |
| Db   | 4444                                   | CTAGACCCACCGTTACTGACTGCTACCTTCCAATTCACCCCAAGAGTTGGGTGCACAG     | 4503        |
| QY   | 4785                                   | atggccctgaagatggaggttctggctgtcgagggagcaagagactctac             | 4832        |
| Db   | 4504                                   | ATTGCCCTGAGGATGGAGGTTCTGGGCTGCGAGGCAAGACCTCTTAC                | 4551        |
| RESULT 7   |  |  |             |
| AR034084   | AR034084                               | 5035 bp  | DNA         |
| LOCUS  | Sequence 1 from patent US 5869292.     | PAT  | 29-SEP-1999 |
| DEFINITION   | AR034084                               |  |             |
| ACCESSION  | AR034084.1                             | GI:5949689   |             |
| VERSION  |  |  |             |
| KEYWORDS   |  |  |             |
| SOURCE   | Unknown.                               |  |             |
| ORGANISM   | Unknown.                               |  |             |
| REFERENCE  | 1 (bases 1 to 5035)                    |  |             |
| AUTHORS  | Voorberg,J.J.                          |  |             |
| TITLE  | Hybrid proteins with modified activity |  |             |
| JOURNAL  | Patent: US 5869292-A 1 09-FEB-1999;    |  |             |
| FEATURES   | Location/Qualifiers                    |  |             |
| source   | 1..5035                                |  |             |
| BASE COUNT   | 1484 a 1127 c 1110 g 1314 t            |  |             |
| ORIGIN   |  |  |             |
| Query Match 47.2%; Score 3748; DB 9; Length 5035;                |  |  |             |
| Best Local Similarity 87.4%; Pred. No. 0;                        |  |  |             |
| Matches 4374; Conservative 0; Mismatches 60; Indels 570; Gaps 2; |  |  |             |
| QY   | 415                                    | ccacatctgcaaatagagctctccacctgctcttctctgtccttctgcatctgtcta      | 474         |
| Db   | 30                                     | CCACCATGGAAATAGAGCTCTCCACCTGCTTCTTCTGTGACCTTTGGCATTCCTGCTT     | 89          |
| QY   | 415                                    | gtgccaccgaagaatactactcctgggtgtcagtygaactgttcataggactatatagc    | 534         |
| Db   | 90                                     | GTGCCACCAAGAAGATACACCTGGGTGAGTGGAAGCTGTCATGGGACATATAGCA        | 149         |
| QY   | 535                                    | atctcgtgtgaagctgctctgtgacgcaagaattccctccatagatgtgcaaatctt      | 594         |
| Db   | 150                                    | ATCTCGGTGAGCTGCTCTGTGACGCAAGATTTCTCTAGAGTGGCAAAATCTTTTCC       | 209         |
| QY   | 595                                    | tcaacacctcagtcgttacaataaagactcgttltgaaatltcaogyltcaoclttca     | 654         |
| Db   | 210                                    | TCAACACCTCAGCTCGTGTACAAAAAGACTCTGTGTGAATTCACGGATCACTTTT        | 269         |
| QY   | 655                                    | acatgctcgaagccaagcgcaacctgtgatgtgtcgtctagtrtctaccatccaggt      | 714         |
| Db   | 270                                    | ACATGCGTAAAGCCAAAGGCCACCTCTGGATGGGTCTCTAGTCTTACCATCCAG         | 329         |
| QY   | 715                                    | ttatcgaatacagtygtctatcaacttaagaaacatlygtcctccatcgtltcagtt      | 774         |
| Db   | 330                                    | TTTATGATACAGATGTGTCATTAACATTAAACATGGCTTCCATCCATCGATCTT         | 389         |
| QY   | 775                                    | ctgtgtgtgtatccctactcgtggaaagcttcttgagggagcgtgaatatgatac        | 834         |
| Db   | 390                                    | CTGTGTGTGTATCCCTACCTGGAAGAGCTTCTGAGGGAGCTGAATATGATAT           | 449         |
| QY   | 835                                    | aaagtgagaagaagatataaagctccctccctcgtgtggaagccatacatatgtc        | 894         |
| Db   | 450                                    | AAAGTGAGAGAAAGATATATAAGCTTCTCTGTGGAGAGCATATATATGTCTGT          | 509         |

|    |      |  |      |
|----|------|--|------|
| OY | 895  | lcccgaaagaaatggtlccaatgggcccctgaaacccactgtgcttactactcaatcatalctt | 954  |
| OY | 896  | lcccgaaagaaatggtlccaatgggcccctgaaacccactgtgcttactactcaatcatalctt | 955  |
| Db | 510  | tccttgaagagaaatgggtccaatggccctctgaccacactgctgcttaccactacatattctt | 569  |
| OY | 955  | cctcatgtgaaccctggttaaaagacttgatctcaggccctcatctggggccctactaagtat  | 1014 |
| Db | 570  | ctcatgtgacacctggtraaagacttgatctcaggccctcatctggggccctactaagtat    | 629  |
| OY | 1015 | gagaaaggagtlctgcccgaaggaaaagacacagacctgtgacaaaatltaactacttltg    | 1074 |
| Db | 630  | gagaaaggagtlctgcccgaaggaaaagacacagacctgtgacaaaatltaactacttltg    | 1075 |
| OY | 1075 | ctgtatttgatgaagggaataaagtgtgcactcagaacaagaactcctgtgtcagaata      | 1134 |
| Db | 690  | ctgtatttgatgaagggaataaagtgtgcactcagaacaagaactcctgtgtcagaata      | 749  |
| OY | 1135 | gggaagtctgaactgtgcccggggccctgggcccctaaatgycacacagtcgaatggtatgaa  | 1194 |
| Db | 750  | gggatctctgcatttgcctcgggcccctgggcccctaaatgycacacagtcgaatggtatgaa  | 809  |
| OY | 1195 | ggctctgccaaggltcgtatctgcatgycacacaggaatacagtlcatctgtgcatgtgaa    | 1254 |
| Db | 810  | ggctctgccaaggltcgtatctgcatgycacacaggaatacagtlcatctgtgcatgtgaa    | 869  |
| OY | 1255 | tgggcaccaactcctgaagtgycactcaaatatctctcgaagtgacacatcttctgtga      | 1314 |
| Db | 870  | tgggcaccaactcctgaagtgycactcaaatatctctcgaagtgacacatcttctgtga      | 929  |
| OY | 1315 | acacatgcgaagggtccttggaatctgcgcacataacttctctactgctcraaaacact      | 1374 |
| Db | 930  | acacatgcgaagggtccttggaatctgcgcacataacttctctactgctcraaaacact      | 989  |
| OY | 1375 | tgatggaccttgagacagtttctactgttctgtcatctctctccacacaaatgatgga       | 1434 |
| Db | 990  | tgatggaccttgagacagtttctactgttctgtcatctctctccacacaaatgatgga       | 1049 |
| OY | 1435 | tggaagactatgtccaaagtgcacagctgctccagaagggaaccccaactcgaatgaaaaa    | 1494 |
| Db | 1050 | tggaagactatgtccaaagtgcacagctgctccagaagggaaccccaactcgaatgaaaaa    | 1109 |
| OY | 1495 | atgaagaacgagaaagactatgcatgatcttacttgatctcgaatgagtggtgcaagt       | 1554 |
| Db | 1110 | atgaagaacgagaaagactatgcatgatcttacttgatctcgaatgagtggtgcaagt       | 1169 |
| OY | 1555 | ttgatgtgacaaactcctccttcccttitaaccaaatctgcgaatgtgccaagaagactc     | 1614 |
| Db | 1170 | ttgatgtgacaaactcctccttcccttitaaccaaatctgcgaatgtgccaagaagactc     | 1229 |
| OY | 1615 | aaacttggttaccttacctgtcgtcgtgaagaaggagactggtgactatgcttccctatg     | 1674 |
| Db | 1230 | aaacttggttaccttacctgtcgtcgtgaagaaggagactggtgactatgcttccctatg     | 1289 |
| OY | 1675 | tcgcccccgatgacagaagttataaagtcgaatttggaacactgggcccctcagaaggat     | 1734 |
| Db | 1290 | tcgcccccgatgacagaagttataaagtcgaatttggaacactgggcccctcagaaggat     | 1349 |
| OY | 1735 | gtagaagatcaaaaaaagtcggaatttaaggcatcacagaatgaacacttlaagactcgt     | 1794 |
| Db | 1350 | gtagaagatcaaaaaaagtcggaatttggaacactgggcccctcagaaggat             | 1409 |
| OY | 1795 | aagcatctcagaatgatacagaagactctgggaccttactcttaagggaagtgtgagaca     | 1854 |
| Db | 1410 | aagcatctcagaatgatacagaagactctgggaccttactcttaagggaagtgtgagaca     | 1469 |
| OY | 1855 | caactgttgatatcatlaaagaatccaagaacagacacataataactcacaactcagaag     | 1914 |
| Db | 1470 | caactgttgatatcatlaaagaatccaagaacagacacataataactcacaactcagaag     | 1529 |
| OY | 1915 | tcactgaatgcgtcccttgtlatccaaggagatbccaaaagtgtaaaaactttgaag        | 1974 |
| Db | 1530 | tcactgaatgcgtcccttgtlatccaaggagatbccaaaagtgtaaaaactttgaag        | 1589 |

[illegible]

|    |      |  |      |
|----|------|--|------|
| Db | 2670 | TCCAATTAGATTAAATGAGAAACTGGGGCAAACTGCAGATCCCTCTTGCTGGGATTAAC        | 2729 |
| OY | 2753 | -----  | 2752 |
| Db | 2730 | ACTATGGTACTCAGATATCCAAAAGAGTGGAAATCCCAAGAGAAATCACCAGAAAAA          | 2789 |
| OY | 2733 | -----  | 2732 |
| Db | 2790 | CAGCTTTTAAGAAAAAGATACCATTTTGTCCCTGAACCTTGTAAGCAATCATGCAA           | 2849 |
| OY | 2753 | -----  | 2752 |
| Db | 2850 | TAGCAGCAATTAATGAGGGACAAAATAATAGCCGGAATACAGCTACCTGGCAAAAGCAAG       | 2909 |
| OY | 2753 | -----cccaacagtccttgaaagcccaacaacggaaa                              | 2784 |
| Db | 2910 | GTAGAGACTGAAGAGCTGTGCTGTCAAAACCCACCACTGTTGAAAGGCATCAAGCGGAAA       | 2969 |
| OY | 2785 | taactcgtactactccttcagtcagatcaagaaggaatctgactatgatatacalatcag       | 2844 |
| Db | 2970 | TAACTCGTACTACTCTTCAGCTCAGATCAAGAGAAATTACATATATGATATACATATCAG       | 3029 |
| OY | 2845 | ctgaaatgaagaagaagatcttgacatataatgaatgaatgaatgaagaagccccgcga        | 2904 |
| Db | 3030 | TTGCAATGAAGAGAGAAATTTTGACATTTATGATGAGATGAATAATCAGACCCCGCA          | 3089 |
| OY | 2905 | gcttccaaagaagaaacagacactatcttatctgctcagctgagagagctcctgagatla       | 2964 |
| Db | 3090 | GCTTTCAAAAGAAACACGACACTATTTATGCTGCAGTGGAGAGGCGTGGAGTTATG           | 3149 |
| OY | 2965 | ggatgagtagctccccaatgctcctaagaacagggctcagagtgagcagtgatccctcagt      | 3024 |
| Db | 3150 | GGATGAGTAGCTCCCGCATGTTCTAAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGT       | 3209 |
| OY | 3025 | tcaagaagaatggtttcccaagaaattacatgaatgagctcttactaagccctataacgctg     | 3084 |
| Db | 3210 | TCAAGAAAGTTGTTTCCAGGAATTTACTGATGAGCTCTTTACTAGGCCCTTATACGCTG        | 3269 |
| OY | 3085 | gagaactaaatgaacattcttggaactccctcggggccatalataaagaacgaagcttgaata    | 3144 |
| Db | 3270 | GAGAACTAAATGAACATTTTGGAGCTCCTGGGGCCATATATAGACAGAGAGTTGAAGATA       | 3329 |
| OY | 3145 | atatcatggttaacttcttgaataatcagggcctctcgtccctatctctatctatctagcccta   | 3204 |
| Db | 3330 | ATATCATGTGTAACCTTTCAGAAATCAGGCGCTTCGTCCCTATTTCTTTATCTTACCTCTTA     | 3389 |
| OY | 3205 | ttctctatggaagaatcacaaggcaaggcaggaagcaaacctagaanaaaccttgcaagcccta   | 3264 |
| Db | 3390 | TTTCTTATGAGGAAGATCAGAGGCAAGGCAAGCAAGAACCTTGAAAAAACTTTGTCAAGCCTA    | 3449 |
| OY | 3265 | atgaaccacaacttactcttggaagtgacaatcataatggcaccacaataaagatgaagt       | 3324 |
| Db | 3450 | ATGAAACCAAACTTACTTTTGGAAAGTCAACATCATATGGCACCCCACTTAAAGATGTAGT      | 3509 |
| OY | 3325 | ttgactgcaaaagcctgagctatctctcgaatgttgaaccttgaaaaaagatgtgaactag      | 3384 |
| Db | 3510 | TTTGACTGCAAAAGCCTGGGCTTATTTCTGTGATGTGACCTGGAATAAATGTGCACTTAG       | 3569 |
| OY | 3385 | gcctgacttggaaccctctcgtctgcacaactaacaacatlgaaacctgtcaatggagac       | 3444 |
| Db | 3570 | GCGTATGTGGACCCCTCTGTCTGTGCACACTATACACATGAAACCTCTATGTGGAGAC         | 3629 |
| OY | 3445 | aagtgaacagtaacaggaatttgctcgtgttttcaaccaatcttgaatgaagcaaaagctggt    | 3504 |
| Db | 3630 | AAGTGAACAGTACAGGAATTTTGCTCTTTTTCACCTTTTATATGAGCAAAAAGCTGCT         | 3689 |
| OY | 3505 | acttcactgaaaaataatgaagaagaacttgcaaggtccctcctgcaatatccaagtgtgaagctc | 3584 |
| Db | 3690 | ACTTCACTGAAAAATATGAGAAAGAACTGCAGGGCTCCTGCAATATCCAGATGGAGATCT       | 3749 |
| OY | 3565 | ccacttttaagaagaaatctcgtcttccatgacatcaatgtgctataatgtgatacaactac     | 3624 |

|    |      |  |      |
|----|------|--|------|
| Dd | 3750 | CCACTTTTAAAGACATTTATTCGCTTCCATGCAATCATGCGTTACTATTAATGGATACACTAC    | 3809 |
| Qy | 3625 | ctggctcttgtaatggtcccaagatccaagaattcgaatgcatctgtctcagcaatggcgacga   | 3684 |
| Dd | 3810 | CTMGCTTAGTAATGGCTCCAGGATCAAAAGGATTGGATGTAATCTGCTCAGCATGGGGACGA     | 3869 |
| Qy | 3685 | atgaanaacatcatctcatctcaatctcaagtggacatggtgtcaactgtacgaaaaaagaag    | 3744 |
| Dd | 3870 | ATGAAACATCATCTATTCTATTCATTCAGTGAGCATGTGTTCCTATGCAAAAAAGAGG         | 3929 |
| Qy | 3745 | agttataaattgacactgtacaactctcatccgggtgtttttgaagcaatggaaatggttac     | 3804 |
| Dd | 3930 | AGTATTAATAATGGCACTCTACAAATCTCTATCCAGAGTGTTTTTATGACACGTGAAAGTGTAC   | 3989 |
| Qy | 3805 | catccaagaactggaattctggcgggtggaaatgaccttaatggcagacatcatagtctggga    | 3864 |
| Dd | 3990 | CATCCAAAGCTGGGAATTTGGCGGGGTGGAAATGCTTATTGGCAGCATCTACATGCTGGGA      | 4049 |
| Qy | 3865 | tgaagcaacttcttcgtgtgtacagaataaaglytcaagctcccttggaaatggtctctg       | 3924 |
| Dd | 4050 | TGAGCACACTTTTCTTGSTGTACAGCAATTAAGTGTGAGACTCCCTGGGAATGCGTTCTG       | 4109 |
| Qy | 3925 | gaacacatagagaattctcaatattacagcttcaggaacatatagacaagtgggccccaaagc    | 3964 |
| Dd | 4110 | GACACATTAGAGATTTTTCACATTACAGCTTCAGGACAAATATGGAAGTGGGGCCCCAACAG     | 4169 |
| Qy | 3985 | tggcgacactcatctatctccgggaatcaatcaatgacctctggagacaaagaagcccccttctt  | 4044 |
| Dd | 4170 | TGGCCACACTTTCATTATTCCGGATTCATCATGATGCTTGAGACCCAGAGGCCCTTTTCTT      | 4229 |
| Qy | 4045 | ggaatcaagglygaatctgttggaaccaatgatatatlaagagcaatcaagaaccaggggtgcc   | 4104 |
| Dd | 4230 | GGATTCAGATGGTGAGCTGTTGGCACCAGATGATTATTACGGGATTCAGAACCCAGGGTGCC     | 4289 |
| Qy | 4105 | gtcagaagctctccagccctctacatctctcagttatcaatcatgatatgctgttgatggga     | 4164 |
| Dd | 4290 | GTCAGAAAGTTCTCCAGGCTCTACATCTCTCAGTTTATCATCTGTAATGTCTTGATGGGA       | 4349 |
| Qy | 4165 | agaagttggagacatctacggagaattccaactggaaaccttaatgltctcttttggcaatg     | 4224 |
| Dd | 4350 | AGAAATGGCACTTATTCGAGGAAATTCACCTGGAACTTAATGAGTCTCTTGGCAATG          | 4409 |
| Qy | 4225 | tgaatcatctgggaatbaaaacacaatattttaaacctccaattatgtctcgatacatcc       | 4284 |
| Dd | 4410 | TGGATTCATCTGGGATTAACACACAAATTTTAAACCTCCATATTTCCTCGATACATACC        | 4469 |
| Qy | 4285 | gtttgcaccacatcatatagatctcgacgaactcttcgaatggagtgtatgggtctgtg        | 4344 |
| Dd | 4470 | GTTTGCACCCCACTCATTAATGAGCTTCGAGACACTTTCGATGGAGTTGATGGGCTGTG        | 4529 |
| Qy | 4345 | atttaaatgaattgagatgacatcttgggaattggagaatgaagaatatcaaatgcaaga       | 4404 |
| Dd | 4530 | ATTTAATAATGTTGAGATGCTCCATTGGGAATGGAGATTAACCAATATCAATGCACAGA        | 4589 |
| Qy | 4405 | ttaactgtctcaatccatctatcaacaatattgttgcacactggtctccctcaaaagctcgac    | 4464 |
| Dd | 4590 | TTACTGCTTCATCTACTTATTACCAATATGTTTGGCCACCTGGTCTCCTTCAAAACCTGCAC     | 4649 |
| Qy | 4465 | ttaacctccaagggaggaatgactgctctggagacactcgaatggatgaataatccaaagaatggc | 4524 |
| Dd | 4650 | TTTCACTTCACAAAGGAGGAGTAATGCTTGAGACACTCAGTGTAATATCCAAAGAGTGGCC      | 4709 |
| Qy | 4525 | tgaagaatggactctccgaagaacaaatgaagaatcaacggaggttaactctcaaggagttaaat  | 4584 |
| Dd | 4710 | TGCAAGTGGACTTCCACAAACACATGAATGAAGTCCACAGAGATTAATCTCAGGAGATTAAT     | 4769 |
| Qy | 4585 | ctctgcactcaacgaatgatatgtaagaagtctcctaactccagagagccaagtgcac         | 4644 |
| Dd | 4770 | CTCTGCTTACCAGCATGTATGTAAGAGAGTTCCTCATCTCCAGCATGTAAAGATGGCCATC      | 4829 |
| Qy | 4645 | aglygaactctcttttttcagaatlgcgaagaatgaagtttttcaaggaaatccaagactctt    | 4704 |
| Dd | 4830 | AGTGAACCTCTTTTTCACAAATGGCAAGTGAAGTTTTTATGGGAATCAAGTCTCT            | 4889 |

QY 4705 tcacacctgtgtgtagactcttagaaccacgcgttactgactgcgtactcgaatccac 4764  
|||||  
Db 4890 TCACACCTGTGTGTAACCTCTTAGACCCACGCTTACTGACTGCTACCTTGAAATTCACC 4949  
|||||  
QY 4765 cccgagatctgggtgtagcaccagatctgcgcgaggaatggaggtctctggcctcgcgagggcagag 4824  
|||||  
Db 4950 CCCGAGATTGGGTGCACACGATTGGCTGAGAGATGGAGGTCTTGCGCTCGAGGACAGG 5009  
|||||  
QY 4825 accctactgactgcgagcagctc 4848  
|||||  
Db 5010 ACCTCTACTGAGGGTGGCCACTGC 5033  
|||||  
RESULT 8  
AR071306 AR071306 5035 bp DNA PAT 18-FEB-2000  
LOCUS Sequence 1 from patent US 5910481.  
DEFINITION  
ACCESSION AR071306  
VERSION AR071306.1 GI:7222194  
KEYWORDS  
SOURCE .  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5035)  
AUTHORS Vooberg,J.J.  
TITLE Hybrid proteins with modified activity  
JOURNAL Patent: US 5910481-A 1 08-JUN-1999;  
FEATURES  
source 1..5035 /organism="unknown"  
BASE COUNT 1484 a 1127 c 1110 g 1314 t  
ORIGIN  
Query Match 47.2% Score 3748: DB 9; Length 5035:  
Best Local Similarity 87.4% Pseq. No. 0;  
Matches 4374; Conservative 0; Mismatches 60; Indels 570; Gaps 2;

Db 510 TCCTGAAGAGAAATGGTCCAAATGGCTCTGACCCACTGTCGCTTACCTATATCTTT 569  
QY 955 cctcatgtgacctggtglaaaagactcgaatlcagcgtccatctggagccctactatgta 1014  
|||||  
Db 570 CTCATGTGACCTTGTAAGAAAGACTTGAATTCAGGCTCTATTGGAGCCCTACTAGATCTA 629  
|||||  
QY 1015 gagaaggagctctggccaaggaagaaacagaccttgcacaaatttactactcttg 1074  
|||||  
Db 630 GAGAAAGGAGTCTGGCCAAAGAAAGACACAGACTTGCACAAATTTACTACTTTTGG 689  
|||||  
QY 1075 ctgtatttgatgaaaggaaggttgcaactcagaaacaaagacccctctgtgagagata 1134  
|||||  
Db 690 CTGTATTGTATGAAGGAAAGTGGCACTCAAGAAACAAAGACTCTTGTATGAGAGATA 749  
|||||  
QY 1135 gggatgctgcatctgctcggtcgtgacctaaatgcacagtcgaatggtatgtaaca 1194  
|||||  
Db 750 GGGATGCTGATCTGCTCGGGCTGCGCTGCTAAATGCACAGCTCAATGATATGTAACA 809  
|||||  
QY 1195 ggtctcgcaggtctgatttgatgtagcacaaggaatcagctctatgcatgtagtggaa 1254  
|||||  
Db 810 GGTCTGTGCCAGGTCTGATTGGATGCCACAGGAATCAGTCTATTGGCATGTGATTGAA 869  
|||||  
QY 1255 tgggacacactcctgaatgtagcactcaatatctcctggaaggtcacacactctgtgagga 1314  
|||||  
Db 870 TGGGACACACTCTCTGAAGTGCATCAATATCTTGAAAGTGCACACATTTCTGTGAGGA 929  
|||||  
QY 1315 accatgcgcagcgctcctctggaatctgcgcaataactctcctactgctcgaacactct 1374  
|||||  
Db 930 ACCATGCGCAGGCGTCTTGGAATCTGCCCATATCTTCTTACTGCTCAAACTCT 989  
|||||  
QY 1375 tgaatgaccttggagacagttcctactgtttgtcatatctctccacaacatgtagtga 1434  
|||||  
Db 990 TGATGACCTTGAGACAGATTCTACTGTTTGTGCATATCTTCCACACCAACATGATGGA 1049  
|||||  
QY 1435 tggaggttatgtcaaatgtagcagctgtccagaggaaccccaactcgcgaatggaataa 1494  
|||||  
Db 1050 TGGAGCTTATGTCAAGATGACAGCTGTCCAGAGGAACCCCACTCGAATGTGAATAAATA 1109  
|||||  
QY 1495 atggaagacggaagactatgtagtactactgctcgaatgtagtgtgtagct 1554  
|||||  
Db 1110 ATGAAGAGCGGAACACATGATGATGATCTTACTGATTCGAATGATGATGTGTCAGGT 1169  
|||||  
QY 1555 ttgatgtgacaactcctcctcttcaaaatcgcctcgaatgtagtgaatgtagtga 1614  
|||||  
Db 1170 TTGATGATGACACACTCTCTTCTTATTCGCAAAATTCGCTGAGTGCACGAAGCATCTTA 1229  
|||||  
QY 1615 aaacttggatcatctacatctgtcgtggaagggagcgtgagctatgctcccttaacc 1674  
|||||  
Db 1230 AAACCTGGGTACATTACATTCGTCTGTAAGAGAGAGAGTGGACTATGCTCCCTTAGTCC 1289  
|||||  
QY 1675 tgcgcccgatgacagaagttataaaagttcaatalttgaaacatgagccctcagcgatg 1734  
|||||  
Db 1290 TCGCCCGGATGACGAAGTTATATAAAGTCAATTTGAAACAATGGCCCTCAGGGAATG 1349  
|||||  
QY 1735 gtaggaagtlacaaaagttccgatttatgcatcacacagatgaaacotttgaagctcgtg 1794  
|||||  
Db 1350 GTAGCAAGTACAAAATAATCCGATTTATGCGATACACAGATGAACCTTTAAAGCTGTG 1409  
|||||  
QY 1795 aagctatcagcatgaaatcagaatccttggagccttactttaggggaagtgtgagaca 1854  
|||||  
Db 1410 AAGCTATTTCAGATGAATCAGGAATCTTGGGACCTTTACTTATAGGGGAATGTGGAACA 1469  
|||||  
QY 1855 cactgtgatlatatlttaagaatcaagcaagcagacacataaacaatacctcaacgaa 1914  
|||||  
Db 1470 CACTGTTGATTATATTTAAGAAATCAGACAGACACCATATTAACATCTACCTTCACGGA 1529  
|||||  
QY 1915 tcaatgtagtccgtccttgtlatltaagggagatlaacaaaaggtglaaaacattggaag 1974  
|||||  
Db 1530 TCACGTGATGCTCGTCTTGTATTCAAGAGATTACCAAAAGGTGTAACATTGGAAG 1589  
|||||  
QY 1975 atttccaaatctgcgcaggaagaataatcaaatataatgtagcaggtgactgtgagaagt 2034  
|||||  
Db 1590 ATTTCCAATTCTGCCAGGAGAAATATTCAATATTAATGACAGTGTAGAAAGATG 1649  
|||||

QY 2035 ggcacaataacagatccctggtgcctgacccgctatacttaagtgttaatatgg 2094  
|||  
Db 1650 GGCACTAATATGATGATCCTGGGCGCCGACCCGCTATTACTTGAATTATATAGG 1709  
QY 2095 agagaagatctagcttcaggaactatggccctctccatctgctacaaagaatctgag 2154  
|||  
Db 1710 AGAGAGATCTAGCTTCAGGAGCTCATTTGGCCTCTCTCATCTGGCTACAAAGAATCTGTAG 1769  
QY 2155 atcaaaagggaaacacagataatgtcagcaagaaggaaatgtcatccgtttctgtatttg 2214  
|||  
Db 1770 ATCAAAAGGGAACCAAGTAATGTGACAGAGAGAAATGATCATCCTGTTTTCTGTATTGG 1829  
QY 2215 atggaacccgaagctgtgacctcaagagaaatatacaagcttctctcccaatccagctg 2274  
|||  
Db 1830 ATGGAACCGGAAGCTGGTACTCTCAGAGAAATATACAAAGCTTTCTCCCAATCAGCTG 1889  
QY 2275 gagtgcagcttgaagatccagaagttccaagccctccaaatcatgacagca tcaatgct 2334  
|||  
Db 1890 GAGTGCAGCTTGAGGATCCAGATGCCAAGGCTCCAAACATCATGACGATCAATGGCT 1949  
QY 2335 atgtttttgataagtttgcagttgtcagtttgttgatgagtggtacatactgtacattc 2394  
|||  
Db 1950 ATGTTTTTGAATAGTTTSCAGTTGTGAGTTGTGATGAGTGCGATACTGTTACATTC 2009  
QY 2395 taagcattggagcagagactgaactcctctctgtctctctctctggaatatacctcaaac 2454  
|||  
Db 2010 TAAGCATTTGAGGACACACATGACTTCTTCTGTCTTCTTCTGATATACCTTCAAC 2069  
QY 2455 acaaaatggtctatlgaaagacacactaccctatccatctcagagaaactgtctca 2514  
|||  
Db 2070 ACAAATGAGTCTATGAGACACACTCACCCATTCCCATTCCTCAGAGAAACTGTCTTCA 2129  
QY 2515 tgtcagatggaaaacccagggtctatgattctgtggtggtccaaactcagactttcgggaa 2574  
|||  
Db 2130 TGTGATGAGAAACCCAGGCTGTATGATTTCTGGGGTCCCAACTCAGACTTTGGGAA 2189  
QY 2575 gaagcagacgccttaactgaagtttctagtttgacaaagacacggtgtattatcag 2634  
|||  
Db 2190 GAGGCATGACCGCTTACTGTAAGGTTTCTACTTGTATTCCAGAGGGGAGAGACGACG 2249  
QY 2635 aggaagctatlgaaatattca-----gcatctgtcgtgtaaaaca 2679  
|||  
Db 2250 ACTATCTGGACCTGGAGAAGATATTCAGTGAAGAGCAGACTCATGACATCGTGACA 2309  
QY 2680 atgccattgaaaccaaagaagttctcccaagaattcaagaacccctcagactcagggaaagc 2739  
|||  
Db 2310 GTCTGATTTGAACCAAGAAGCTTCTCCAGAAATTCAGACCCCTAAGCATAAGCCAAAAGC 2369  
QY 2740 aatttaatgcca----- 2752  
|||  
Db 2370 AATTATATGCGACACACAAATTCAGAAATGACATAGAGAACTGACCCCTGTTTGAC 2429  
QY 2753 ----- 2752  
Db 2430 ACAGAACACCTATGCTTAATAATACAAATGTCTCCTAGTATGTTGTATGCTTGGC 2489  
QY 2753 ----- 2752  
Db 2490 GACAGAGTCTACTCCACATGGGCTATCTTATCTGATCTCCAGAGGCCAAATATGAGA 2549  
QY 2753 ----- 2752  
Db 2550 CTTTTCATGATGATCCATCACTGAGACATAGACATATATACAGCCTGTCTGAATGA 2609  
QY 2753 ----- 2752  
Db 2610 CACACTTCAGGCCACAGCTCCATCACAAGTGAGGACATGTAATTTACCCCTGAGTCAAGCC 2669  
QY 2753 ----- 2752  
Db 2670 TCCATTTAAGATTAATAGAAACCTGGGGACAACTGCAGATCTCTTGTGCTTGGGATTAAC 2729

QY 2753 ----- 2752  
Db 2730 ACTATGTACTCAGATACCAAAAGAGAGTGGAATCCCAAGACAGATGCCAGAAAAA 2789  
QY 2753 ----- 2752  
Db 2790 CAGCTTTAAGAAAAAGGATACCATTTTGTCCCTGAACGCTTGTAAGCAATCATGCA 2849  
QY 2753 ----- 2752  
Db 2850 TAGCAGCAATTAATGAGGACCAAAATAAGCCGAATTAAGATCACTGGGCAAAAGCAG 2909  
QY 2753 -----ccacagacttgaacgcgcatacaaggga 2784  
|||  
Db 2910 GTAGACTGAAAGGCTGCTGCTCAAAAACCCACAGCTTTTAACGCCATGAAAGGGGAA 2969  
QY 2785 taactgtactactctcttaagtcagatccaagaagaattgactatgatatcatalcaag 2844  
|||  
Db 2970 TAACTCGTACTACTTCTCAGTACAGATCAAGAGAAATGACTATGATGATATCATATCAG 3029  
QY 2845 ttgaaatgaaagaaggaatttgacattatgatatgagatgaaatcagaagcccgca 2904  
|||  
Db 3030 TTGAATGAAGAGAGAAATTTGACATTTATGATGAGGATGAATAATCAGAGCCGCCCA 3089  
QY 2905 gcttcaaaagaacacagacactatttatgtctgcagtgagaagctcgggatatg 2964  
|||  
Db 3090 GCTTTCAAAAGAAAACACGACACTATTTTATTGCTGAGTGAGAGGCTTGGAATTAAG 3149  
QY 2965 ggaatgaatgctccccaatgtctaaagaaacagggctcagaagtgcagtgctccctcag 3024  
|||  
Db 3150 GGATGAGTAGCTCCCAACATGTCTTAAGAAACAGGGCTCAGAGTGGCAGTCCCTCAGT 3209  
QY 3025 tcaagaagaagtgtttccagaagaatttactgtagtgcctcttaaccagccctatacgtg 3084  
|||  
Db 3210 TCAAGAAAGTGTGTTTCCAGGAATTTACTGATGAGTGGCTCCTTTAATCAGGCTTAATACCGTG 3269  
QY 3085 gagaacaaatlgaaacatttggagactccgtggggccatatataagaagcagaattgaagta 3144  
|||  
Db 3270 GAGAACATAATGAACATTTTGGGACTCCTGGGGCCATATATAGACAGCAAGTTGAAGTA 3329  
QY 3145 atatcatgtaacttcaagaatcaggcctcgtccctatctctctatctatctagcctta 3204  
|||  
Db 3330 ATATCATGTGTAACCTTCAGAAATCAGGCTCTGCTCCCTATTCCTTCAATCTAGCTTA 3389  
QY 3205 tttcttatggagaagatcagaaggcagaaggaagaaactcagaanaaacttgtcaagccta 3284  
|||  
Db 3390 TTTCTTATGAGAGAGATCAGAGGCAAGGACGAACTTAGAAAAACTTGTCAAGCCTTA 3449  
QY 3265 atgaacccaaacttacttcttggaaagtgcacatcatatgycaccccaatgaagatggt 3324  
|||  
Db 3450 ATGAACCAAAACTTACTTTTGGAAAGTGCACATCATATG6CACCCATTAAGATGAGT 3509  
QY 3325 ttgactgcacaaagcctgggctatctctctgtatgttgcaccttggaaagaatgtgcactcag 3384  
|||  
Db 3510 TTGACTGCAAAAGCCTGGGCTTATTTCTGTATGTTGACCTGGAAAAAGATGCACTCAG 3569  
QY 3385 gccctgattggaacccctctctgtctgcacacactaacacacgtgaacccctgtcattggaagac 3444  
|||  
Db 3570 GCCCTGATTGACCCCTTCTGCTGTGACACATAACACTGAACCCCTGCTCATATGGAGAC 3629  
QY 3445 aagtgcagtagaagaaatttgctctgttcttccacacacttggagagacaaagcgtg 3504  
|||  
Db 3630 AAGTGACAGTACAGGAATTTGCTGTGTTTTCACCATCTTTGATGAGAACCAAAAGCTGCT 3689  
QY 3505 acttcaactgaaatatalgaaagaacatgcagggctccctcgaatataccagatggaagatc 3564  
|||  
Db 3690 ACTTCACTGAAAAATATGAAAGAACTGACAGGGCTCCCTCAATATCCAGATGGAAGATC 3749  
QY 3565 coactttlaaaggaatatcgccttcaatcaatcaatgagctcagatatgataactac 3624  
|||  
Db 3750 CCACTTTTAAAGAGAAATTAATGCTTCATGCAATCAATGAGCTTCAATATGATGATACACTAC 3809  
QY 3625 ctggtctagtaatggtcagatcaagaatctcagatglatctgctcagatggtgcga 3684

|   |      |  |       |
|---|------|--|-------|
| D | 3810 | CTGGCTTAGTAAAGGCTCAGGATCAAAAGATTGATGATGCTCTACGATGGGACACA         | 3865  |
| O | 3685 | atgaaacaacatcattcattcattcagctgagacatgcttcaactgtaacgaataaagagg    | 3744  |
| D | 3870 | ATGAANAACATTCATTTCATTTCAGTGGACATGTTGTCACGTGACGAAAAAAGAGG         | 39229 |
| O | 3745 | agtataaattgcaactgtacaaactctacacaggctgtttttgagacagtggaaatgltac    | 3804  |
| D | 3930 | AGTATATAAAGGCACTGTACAACTCTCATTCACAGGTGTTTGGAAACAGTGGAAATGTATAC   | 3985  |
| O | 3805 | catcacaagtgtgaatttggcgggtggaaagccatttggagagcatctacatgtgtgga      | 3866  |
| D | 3990 | CATCCAAAGCTGGAAATTTGGGGGTGAATGCCATTATTTGGGACATCTCAATGCTGGGA      | 4049  |
| O | 3865 | tggagcaactttttctgtgtgtacagcaataagtttcaagaccctctggaaatggcttctg    | 3924  |
| D | 4050 | TGAGCACACTTTTCTGCTGTGACGACATATAGTGTACAGCTCCCTGGGAATGGCTTCTG      | 4109  |
| O | 3925 | gacacattagaagatttccagatliaacgcttccaggacaataltgacaagtggcccccaagc  | 3984  |
| D | 4110 | GACACATTAGAGATTTTCAGATTACACTTTCAGACATATGAGACAGTGGGCCCCCAAGC      | 4169  |
| O | 3985 | tggccagacttaattatccggtatcaatccaatgtcctgtgagacccaaggagccctttctt   | 4044  |
| D | 4170 | TGGCCAGACTTCATTATTTCCGGATCAATCAATGCTTGGAGCACAAGAGGCCCTTTCTT      | 4229  |
| O | 4045 | ggaatcaagtggaatctgttggacaacatgtatttcaagggatcaagaccagggtgcc       | 4104  |
| D | 4230 | GGATCAAGTGTGATCTGTTGGCACCAATGATATTACAGCGCATCAAGACCCAGGCTGCC      | 4289  |
| O | 4105 | gtcagaagttcccaagccttacaactctccagtttcaatcaatglatagttctgtatggga    | 4164  |
| D | 4290 | GTCAGAAGTCTCCAGCGCTTACATCTCTCAGTTATCATCATGATATATGTTATGGGA        | 4349  |
| O | 4165 | agaagtggcagacttattcggaggaaattccaactgtgaaccttaatgtgtctcttggcaatg  | 4224  |
| D | 4350 | AGAAAGTGGAGACTTATTCGAGGAAATTCACATGGAACTTATATGTCCTTTGGCAATG       | 4409  |
| O | 4225 | tggattcatctgggataaacaacataattttaaaccctccaattatgtctcgatcacatcc    | 4284  |
| D | 4410 | TGCAATTCATCTGGGATTAACACACATATTTTAACTCCCAATTAATGCTCGATACATCC      | 4469  |
| O | 4285 | gtttgacccaactcattatagaacttccgacacacttccgaatggagttgaatggctgtg     | 4344  |
| D | 4470 | GTTTGACCCCAACATATTATAGCTTTGCGACACTCTTCGCATGGAGTTGATGGGCTGTG      | 4529  |
| O | 4345 | atttaaatgtttgcagatgccattgtggaatggaggtaaagcaatatcagatgcacaga      | 4404  |
| D | 4530 | ATTTTAATATATCTTCAGCATCCCATTTGGGAATGGAGAGTAAAGCAATACATCCACAGA     | 4589  |
| O | 4405 | ttaacttcaactacttattacaacatatgtttgcaactgtgtctcttcaaaagctgcac      | 4464  |
| D | 4590 | TTACTGTCTTACTCTACTTAAACATATGTTTTGCCACTGTCTTCCTTCAAAAGCTGCAC      | 4649  |
| O | 4465 | ttaacttccaaaggagagtaatgtccttggagaccttcaagttgaataatccaaagagtgc    | 4524  |
| D | 4650 | TTTCACTCCAAAGGAGAGATTAATGCTCGAGACCTCAGTGAATTAATCCAAAGAGTGGC      | 4709  |
| O | 4525 | tgcgaatggaacttccagaagaacaaatgaaagtcaaggaatfaactcaactcaggagataaat | 4584  |
| D | 4710 | TGCAAGTGGACTTCCAGAAACATATGAAGTCAAGAGATTAACATCAAGGAGATAAAT        | 4769  |
| O | 4585 | ctctgtcttaccagaatgtatgttaagaagattcccatctccagaagaatcaaatggccatc   | 4644  |
| D | 4770 | CTTGTCTTACCAAGATATATGTAAAGAGATTCTCATCTCCAGCATCAAGATGGCCATC       | 4829  |
| O | 4645 | agtggaatcctcttttccagaatgtgcaaaatgaaagtttccagaaggaaatcaaaagctctc  | 4704  |
| D | 4830 | AGTGGACTCTCTTTTTCAGAAATGCAAAAGTAAAGTTTTTCAGGGAATCAAGACTCTCT      | 4889  |
| O | 4705 | tcaacactgtgtgtaactctctagaccacogttactgtaactcgtaacttcaatcaacc      | 4764  |

|            |                                      |   |      |             |
|------------|--------------------------------------|---|------|-------------|
| Db         | 4890                                 | TCACACCTGTGTGTCGATACCTCTTAACACCCACCGTTACTGACTGCTACCTTCGAAATTCAC | 4943 |             |
| Qy         | 4765                                 | cccaagatctgggtgacaccagaatgccttgagatlgagagttcttggcttgcagagacag   | 4824 |             |
|            |                                      |   |      |             |
|            |                                      |   |      |             |
| Db         | 4950                                 | CCGAGAGTTGGGTTGACACAGATTCCCTGAGGATGAGAGTTTGGGCTCGAGGCACAG       | 5009 |             |
| Qy         | 4825                                 | accttactgactgcagcagttc  | 4848 |             |
|            |                                      |   |      |             |
|            |                                      |   |      |             |
| Db         | 5010                                 | ACCTTACTGAGGGTGCGCCACTGC  | 5033 |             |
| RESULT     | 9                                    |   |      |             |
| AR029098   |                                      |   |      |             |
| LOCUS      | AR029098                             | 4334 bp   | DNA  | PAT         |
| DEFINITION | Sequence 38 from patent US 5859204.  |   |      | 29-SEP-1999 |
| ACCESSION  | AR029098                             |   |      |             |
| VERSION    | AR029098.1                           | GI:5941071  |      |             |
| KEYWORDS   |                                      |   |      |             |
| SOURCE     | Unknown.                             |   |      |             |
| ORGANISM   | Unknown.                             |   |      |             |
| REFERENCE  | Unclassified.                        |   |      |             |
| AUTHORS    | 1 (bases 1 to 4334)                  |   |      |             |
| TITLE      | Lollar,J.S.                          |   |      |             |
| JOURNAL    | Modified factor VIII                 |   |      |             |
| FEATURES   | Patent: US 5859204-A 38 12-JAN-1999; |   |      |             |
|            | Location/Qualifiers                  |   |      |             |
|            | 1..4334                              |   |      |             |
|            | /organism="unknown"                  |   |      |             |
| BASE COUNT | 1109 a 1126 c 1085 g                 | 1014 t  |      |             |
| ORIGIN     |                                      |   |      |             |

[illegible]



|    |      |  |      |
|----|------|--|------|
| Db | 543  | CACGGACCGCTGGGAAGACTGAAATTCGGGGCTCATTTGGACCCCTGGTTGTAGAA       | 602  |
| Oy | 1017 | gaaggaaatctgcgcgaagaaagaaacacagaccttgcaaaattactactcttgcct      | 1076 |
| Db | 603  | GAAGGAGATCTGACACAGAAAGAACCCCAAACTGCAAGAAATTTTACTACTTTTGGT      | 662  |
| Oy | 1077 | gtatttgatgaaggaaaaglttgccatcagaacaagaactccttga-tgcagatag       | 1136 |
| Db | 663  | GCTTTTGATGAAAGGAAAAATTTGGCACTCAGCAAGAAATGACTCTCGACAGGGCCATG    | 722  |
| Oy | 1137 | gatctgatctgtctgcggccctgaattcaaatgcaacagtgaaatggttatgaaacg      | 1196 |
| Db | 723  | GATCCCGCACTGCGAGGGGCCAGCCCTGCATACACAGTAAATGCTATGTCAACAGG       | 782  |
| Oy | 1197 | tctctgcgaagtgctgatctgatccacaagaatacagctcatatggatgtga-ttggatg   | 1256 |
| Db | 783  | TCTTGCCAGGTCGTGATCGATGTCATAGAAATCAGTCTACTGGCAGCTGATTTGGAAATG   | 842  |
| Oy | 1257 | ggcaccaactccttgaaagtgcactaatalctcctgaaagttcacacatlttcttgagaac  | 1316 |
| Db | 843  | GGCACACAGCCCGGAAGTGCACTCCCATTTTCTTGAAGGGCACACGTTTCTCGTAGGCAC   | 902  |
| Oy | 1317 | catgcgcagcgctctcttggaactctgcgcataacttccctactgcttcaaacactctg    | 1376 |
| Db | 903  | CATGCGCAGGGTCTTCCTTGGGATCTCGCCACTCAACTTCTCTACGTCAGCAATTTCCG    | 962  |
| Oy | 1377 | atggaaccttgacaagttctactggttttgcatactctctccacaagaatgagacg       | 1436 |
| Db | 963  | ATGGAACCTTGCCAGTCTTCACGTGTTTGTGCATATCTTCCACACCATGATGTGGCATG    | 1022 |
| Oy | 1437 | gaagcttatgtcacaagatagacagctgtccagaaagaaacccaactcagaatgaaataat  | 1496 |
| Db | 1023 | GAGGCTACGTCAGAGATGAAAGCTCGCGCGAGAGACCCCACTGGAGGAAACCTGAT       | 1082 |
| Oy | 1497 | gaagaacgcgaagaagactatgatgtatctactgattctgaaatggaatggtgcagttt    | 1556 |
| Db | 1083 | GAAGA---GGAAAGTATGTAGTGAACAATTTGTACAGCTCGCATGTGACGTGCTCGGCTC   | 1139 |
| Oy | 1557 | gabatgaacaactcctcctcttcaatccaaatcgctagtttgccaagaagaatctctaa    | 1616 |
| Db | 1140 | GATGCTGACGAGTGTCTCCCTTATTCCAATCGCTCGGTTGCCAAGAAAGATCCCAAA      | 1199 |
| Oy | 1617 | acttggttacattacatctgtctgtcgaagaaagagacttggaactatgctccttagcttc  | 1676 |
| Db | 1200 | ACGTGGGTGCACATCACTCTGCAAGAGAGAGAGACTGGGACTACGCCCGCGGTCCC       | 1255 |
| Oy | 1677 | gccccgaatgacagaagtataaagltcaatttgaacaatgyccttcacgcgaat         | 1736 |
| Db | 1260 | AGCCCCAGTGCAGAAAGTTATTAAGTCTCTACTTGAACAGTGGTCTCAGCAATTTGGT     | 1319 |
| Oy | 1737 | aggaagaatcaaaaaagctccgaattatggtcatcacagatgaaactcttaagaactcgtaa | 1796 |
| Db | 1320 | AGGAATATCAAAAAAGCTGCATTCGTCTTACACGAGATGTAACTTTAAAGCTGTGTAA     | 1379 |
| Oy | 1797 | gcta-ttcagcatgataagaaactctggaaccttcttca-tgggaagttggaagaca      | 1856 |
| Db | 1380 | GCTATTCGATGATGATAGGAATCTCTGGGACCTTTACTTATGGAGAACTTGGAGACACA    | 1439 |
| Oy | 1857 | ctgttgatattattaagaa-tcaagcaagcagacataatacatctacccctcaagaa-tc   | 1916 |
| Db | 1440 | CTTTTGATTTATTTAAGATTAAGCAAGCAGCCGCACTTATTAACATCCATCCATGTGATTC  | 1499 |
| Oy | 1917 | actgatctgcgtctcttgat-tcaaggaatataccaaaagtgataaacatttgaagat     | 1976 |
| Db | 1500 | ACTGATGTCAAGCGTTTGGCACCCAGGAGAACTTTTAAAGTTTGAAGAACATTTTGAAGAC  | 1559 |
| Oy | 1977 | tttcaaatctgcgcaggaagaatcatccaataaa-tgagagtgacgttgagaagatgg     | 2036 |
| Db | 1560 | ATGCGCATTTCTGCCAGAGAGACTTTCAAGTATTAATGACACGTACTGTGGAGATGGG     | 1619 |
| Oy | 2037 | ccaactaatcagatccctcggtgcctgaccgcgtatctactagtttcgttaataatgag    | 2096 |

|    |      |  |      |
|----|------|--|------|
| Db | 1620 | CCACCAAGTCCGATCCTCGGTGCTGACCCCGCTACTACTCGAGCTCCATTAACTAGAG         | 1679 |
| Qy | 2097 | agaagtcacgtctcaaggactcaatgagccctctccatctctgatacaagaatctgtagat      | 2156 |
| Db | 1680 | AAAGATCTGGGCTTGGGACCTCATTTGGGCCCTCTCTCATCTGCTACAAAGAACTGTAAAC      | 1739 |
| Qy | 2157 | caaaagaggaacacagataatgltcagacaagaagaaatgltcatcctglttcgtlatltagat   | 2216 |
| Db | 1740 | CAAAAGGAGAAACCGAGTATGATGTCCAGCAAAAGAAACGTCATCTCGTTTCTGTATTTCAT     | 1799 |
| Qy | 2217 | gagaaacgaagctggtactccctcaagaggaataatacaacgcttctctcccaatcagctgga    | 2276 |
| Db | 1800 | GAGAAATCAAACTCTGTACTCTTCGCGAGAAATATTTACGCGTCTCTCCCAATCCGATGGA      | 1859 |
| Qy | 2277 | gtcagactgtgagatgccagagatgccagagctcccaacatgatcagacgaatcaatgtgcat    | 2336 |
| Db | 1860 | TTACAGCCCCCAGATCCAGAGATTCCAAAGCTTCTAAACATATGACACAGCATCAATGGCTAT    | 1919 |
| Qy | 2337 | gltttgtagatgttcagactgtgcagttgttctgttcagatggtgacgaactgtgatactcta    | 2396 |
| Db | 1920 | GTTTTTGATACCTTGACAGCTGTGTGGTGTGTTTGTCACAGAGGTGGACATCGTACATTCCTA    | 1979 |
| Qy | 2397 | agcaatgtagacacagactgtaactccctctctgcttctctctctctctctgataactcctaacac | 2456 |
| Db | 1980 | AGTGTGGAGACACAGACGGACTTCCTCTCCGTCCTCTCTCGGTACACCTTCCAAACAC         | 2039 |
| Qy | 2457 | aaataggtctatgtagaac  | 2516 |
| Db | 2040 | AAATATGCTTATGTAAGACACACTCCCTGTTCCCTCTCAGGAGAAACGGCTCTTCATG         | 2099 |
| Qy | 2517 | tcgaatggaacaaacccaaggtctctatgtaattctgtaggtgcacacacacacacacacacacac | 2576 |
| Db | 2100 | TCAATGAGAAACCCAGGCTCTGGGTCCTTAGGGTGCACACACTCAGACTTGGGAGAACGA       | 2159 |
| Qy | 2577 | ggcagtgaccgccttactgtaagagttctagtgtgtgacagaacacacgtgtatattatacag    | 2636 |
| Db | 2160 | GGGATGACACGCTTACCTGAAAGGTGTATAGTTGTGACAGGACATTTGCTGATTTATATAC      | 2219 |
| Qy | 2637 | gacagttatggaagatatttcaagctactgtgtgagtaaaacaatgacctgtaaccaaga       | 2696 |
| Db | 2220 | AACACTTATGAGATATTCCAGGCTTCTTGCTGTAGTGGAAAGATGTCAATTGAAACCCAGA      | 2279 |
| Qy | 2697 | agcttctcccgaaatltcaagaacacctlagactlagcaaaagcaattaatgtgccaccca      | 2756 |
| Db | 2280 | -----  | 2279 |
| Qy | 2757 | ccagtccttgaaacgcacataacggaataaactcgtactactctctcagtcagatacaag       | 2816 |
| Db | 2280 | -----GACATTAAGCCTTCTTACCGCGGAGGAAAC                                | 2315 |
| Qy | 2817 | gaaatgactatgtagtatacatatacgtttgaaatgaaatgaaagaattttgacattat        | 2876 |
| Db | 2316 | AAAAATGCAATGATGATATCTTCTCAACTGAACCAAGGAGAAAGATTTGTACATTTTAC        | 2375 |
| Qy | 2877 | gatagatgtgaaatcaacagcccccccgagcttcaaaagaaacacacgacataattat         | 2936 |
| Db | 2376 | GGTAGAGATGAATATGACAGACCTCCGACGTTTCAGAAAGAACCCGACACTATTTCATT        | 2435 |
| Qy | 2937 | gtcgcagtggaagaggtctcgtggaatlatggaatgtagtagtctcccaacatgltctaaagaac  | 2996 |
| Db | 2436 | GCTCGGAGGACGAGCTCTGGGATTTACGGGATGAGCGAATCCCCCGGGGCGCTAAAGAAC       | 2495 |
| Qy | 2997 | aggagctaaagtgtagcagtgctccctcaagttccaagaaagtgttttccaagaattacgat     | 3056 |
| Db | 2496 | AGGGCTCAGAACGGAGAGGTGCTCGGTTTCAGAAAGTGTCTTCCGGGAAATTTGTCTAC        | 2555 |
| Qy | 3057 | ggcccttctactcagccctataacgttggaagaaatgaatgaacatttggaactcctgag       | 3116 |
| Db | 2556 | GGCTCTTTCACAGCGCGCTGTACCCGGGGAATCTCAACAAACACTTGGGCTCTTGGA          | 2615 |
| Qy | 3117 | ccatataaagagcagaaglttgaagataataatcatgtltaactttcagaatcagagcctct     | 3176 |
| Db | 2616 | CCCTACATTCAGAGGGAAGTTGAAGCAACATCATGTGTAACTTTTAAAAACCGAGCGCT        | 2675 |



|    |      |   |       |
|----|------|---|-------|
| QY | 3177 | cgfcccattccctctatctctagcctattctctatgaagaaatcgaagcaagagca          | 3235b |
|    |      |   |       |
| Db | 2676 | CGCCCTATTCTTTACTCTGAGGCTTATTTCTTATCCGGATGATCGAGCAAGGGGCA          | 2735b |
|    |      |   |       |
| QY | 3237 | gaaccagaaaaaactctgcaagcccaatgaaaccaaacacttactcttgaaagtgcga        | 3296b |
|    |      |   |       |
| Db | 2736 | GAACCTCGACACAACTTCGTCACGCGCAATGAACCAAGAACTTCTTTGGAAGTGAG          | 2795b |
|    |      |   |       |
| QY | 3297 | catcatatgagcacccactaaagaatgagtttgactgcgaagccctggtcttactgat        | 3356b |
|    |      |   |       |
| Db | 2796 | CATCATGTGGCACCCACAGAAAGAGATTTTGACTGCAAAAGCCTGGGCTACTTTTCGAT       | 2855b |
|    |      |   |       |
| QY | 3357 | gttgacctggaaaaaataatgagcactaagcgctgatgtgacctctgtgtgcacaact        | 3416b |
|    |      |   |       |
| Db | 2856 | GTTGACCTGGAAAAAGATGTGCCTCAGGCTTGATCGGCCCTTGTGATGTGCGCGCC          | 2915b |
|    |      |   |       |
| QY | 3417 | aaacacatgaaccctctctcatggagaagaatgacatcacagaatctgctcttctc          | 3476b |
|    |      |   |       |
| Db | 2916 | AACACCTGAAGCCTGCTACGAGTAGCAAGTAGACCGTCGAGAAGATTTGCTCTGTTTTTC      | 2975b |
|    |      |   |       |
| QY | 3477 | accatctctgaatgagaccaaagctgttacttcaactcgaataatgysaaagaactcgaag     | 3536b |
|    |      |   |       |
| Db | 2976 | ACTATTTTGTGATGAGCAAAAGACTGTGATCTCTGTAATAATGGAAGAAAGAACTGCCGG      | 3035b |
|    |      |   |       |
| QY | 3537 | gtccctcgtgaatatccagaatgagaatcccaactttaaagaaatttgcgtccatgca        | 3596b |
|    |      |   |       |
| Db | 3036 | GGCCCTCGCACCTGCAAGATGAGAGACCCACTCTGAAAGAAACTATGCTTCATGCA          | 3095b |
|    |      |   |       |
| QY | 3597 | atcaatgctgctacaataatgatacactaccctggtcttaataatgctcaagaatcaagaatt   | 3656b |
|    |      |   |       |
| Db | 3096 | ATCAATGGGCTATGTGATGGATACACTCCCGGCTATTATAGCTCGAAGTCAAGATC          | 3155b |
|    |      |   |       |
| QY | 3657 | cgatggtatctgctcagcatggtgcaagcaatgaaacatccatctatcatctcagtga        | 3716b |
|    |      |   |       |
| Db | 3156 | CGATGTGATCTGCTCGACATGAGGCGAGCAAGTAAATATCCATTGATTCATTTAGCGGA       | 3215b |
|    |      |   |       |
| QY | 3717 | catgtgtctacgttaagaaaaaagaagagatataaaatggcagctgacatctctatcca       | 3776b |
|    |      |   |       |
| Db | 3216 | CACGTTCATGTTACGTACGGAAAAAGAGAGAAATATTAATATGGCCGCTACAAATCTTATCCG   | 3275b |
|    |      |   |       |
| QY | 3777 | ggtgctctctgaagacagtgysaaatgttaccatccaaagcttgaatttggcgggtgaaatgc   | 3836b |
|    |      |   |       |
| Db | 3276 | GGTGTCTTTGAGACAGAGGAAATGCTATCCGTCCAAAGTTGGAATTTGGCAATGAAATGC      | 3335b |
|    |      |   |       |
| QY | 3837 | cttatctggcgaagacatctacaatgctggtggaatgaagcaacttctctgtgtacagcaataag | 3896b |
|    |      |   |       |
| Db | 3336 | CTGATTTGGCGAGACCTGTGCAAGTGGGATGAGCAGCACTTCTCGGTACGACGAAGAG        | 3395b |
|    |      |   |       |
| QY | 3897 | gtctgaagcccccctgggaatgtgtctcgtgagacacatagagattctgaatttaagctctca   | 3956b |
|    |      |   |       |
| Db | 3396 | TGTCAGGCTCCACTGGGAATGCTTCTGTGGACGCTTAGAAGATTTTTCAGATCACAGCTTCA    | 3455b |
|    |      |   |       |
| QY | 3957 | ggacaaatgysaagatgysgcccccaaaagctctgcccagactcatatctccgagatcaat     | 4016b |
|    |      |   |       |
| Db | 3456 | GGACAGATGAGACAGTGGGCCCAAAGCTGGCCAGACTTCATTATTCGGATCAATCAAT        | 3515b |
|    |      |   |       |
| QY | 4017 | gacctgagacccaagaagacccttctctcttggaatcaaggtgatcgtgttgccaacatgact   | 4076b |
|    |      |   |       |
| Db | 3516 | GGCTGGAGACCAAGAGATCCCCACTCCTGTGATCAAGGTGATCTGTTGGCACCAATGATC      | 3575b |
|    |      |   |       |
| QY | 4077 | attcaagcatcaaaagacccaaggtgtgcccgtccaaagtctccagcctctcaatctccag     | 4136b |
|    |      |   |       |
| Db | 3576 | ATTCAAGCATCATGACCCAGGGGTGCCGTGACAGATTTTTCAGCCCTTACATCTCCAG        | 3635b |
|    |      |   |       |
| QY | 4137 | tttatcatcatatagctcttgaatggaaagaatgysagactctacaggaatctccact        | 4196b |
|    |      |   |       |
| Db | 3636 | TTTATCATCATGTACAGTCTTTGACGGGAGGAACTGGCAGAGTTTACCGAGGAATTTCCAG     | 3695b |
|    |      |   |       |
| QY | 4197 | ggaaccttaatggtctctcttggcaatgysatcatctcgtggaataaaacaatatcttt       | 4256b |
|    |      |   |       |
| Db | 3696 | GGCACCTTATATGGTCTTCTTTGGCAATGAGAGCACTGTGGGATTTAAACCAATATTTT       | 3755b |

|    |      |   |      |
|----|------|---|------|
| QY | 4257 | aaactccaatatatgttcgataacatccgtttgaccccaactcaatataagatcgcagc     | 4310 |
| Db | 3736 | AACCTCCGATTGGGCTCGGTACATCCGTTTCCACCCAAACATTACGCAATCCGCGAC       | 3815 |
| QY | 4317 | actctcgcacatgagatctgcatgctgctgattcaatagttcagcatgcatcttggaatg    | 4376 |
| Db | 3816 | ACTCTGCCATGGAGATTGATGCGCTGGATTTTAAACAGTTGGCAGCATGCCCTGGGAATG    | 3875 |
| QY | 4377 | gagagtaagaacatatcatagatgacagatctactgtcttcaaccttcaactaataatgltt  | 4436 |
| Db | 3876 | CAGATTAAACGGATATCAGACTCACAGATCACGGCTCCCTCCACCTAAAGCAATATATT     | 3935 |
| QY | 4437 | gccacctggtctccttcaaaagctcgacttcaactcctccaaggagaagatgctcgcgaga   | 4496 |
| Db | 3936 | GCCACCTGGTCTCCTTCCACAAAGCCCGCACTTCACTCCAGGGCGGACGAAATCCGCGGA    | 3995 |
| QY | 4497 | cctcagatgtaaataccaagaagctgctcgaatgagcttcagaagaacatgaagctc       | 4556 |
| Db | 3996 | CCCCGGGTGACACACGCAGAGAGATGGCTCAGTGGACTCTGCAGAAACGCTGAAGTTC      | 4055 |
| QY | 4557 | acagagtaacttaactcagaggagtaaaatctctgcttaccagaatglatgtgaagatlc    | 4616 |
| Db | 4056 | ACAGGCAATCACACACCCAGGGCGTGAAGTCCCTGCTCAGCAGCATGATGTGAAGATTC     | 4115 |
| QY | 4617 | ctcactctccagcagtcagaatgacctcaagtactctcttttttcagaatgycagaagta    | 4676 |
| Db | 4116 | CTCGTGTCCATATCTACGAGCGCGCGCGCTGGACCCCTTTTCTTCAGACGCGCACACG      | 4175 |
| QY | 4677 | aagcttttcaagggaatacaagactccttcaaccctgtgtgaactctctagaccacccg     | 4736 |
| Db | 4176 | AAGGTTTTTCAGGGCAATTCAGACTCTCTCACCCCGCTGTGAAGCTCTGACCCCGCG       | 4235 |
| QY | 4737 | ttactgactcgtctacaccttcgaattcaaccaccagatcttggttgaccagatlgccctgag | 4796 |
| Db | 4236 | CTGTTCACGGCCTACGTAAGGATTCACACCCAGCAGCTGGGGCGACACACATGGCCTGAGG   | 4295 |
| QY | 4797 | atggaagttctggcctgcgcaggaacagagactctactga                        | 4835 |
| Db | 4296 | CTCGAGGTTCTAGAGATCTGAGCGACAGGAGATCTCTACTGA                      | 4334 |

|           |            |   |            |                  |
|-----------|------------|---|------------|------------------|
|           | RESULT     | 10  |            |                  |
| CYU089927 | LOCUS      |   |            |                  |
|           | DEFINITION | CYU089927   | 6786 bp    | DNA circular SYN |
|           | ACCESSION  | U89927  |            |                  |
|           | VERSION    | U89927.1  | GI:2190694 |                  |
|           | KEYWORDS   | .   |            |                  |
|           | SOURCE     | unidentified cloning vector.  |            |                  |
|           | ORGANISM   | artificial sequence; vectors.   |            |                  |
|           | REFERENCE  | 1 (bases 1 to 6786)   |            |                  |
|           | AUTHORS    | Holtz,A. and Lou,Y.   |            |                  |
|           | TITLE      | PHIST1, complete sequence   |            |                  |
|           | JOURNAL    | unpublished   |            |                  |
|           | REFERENCE  | 2 (bases 1 to 6786)   |            |                  |
|           | AUTHORS    | Holtz,A. and Lou,Y.   |            |                  |
|           | TITLE      | Direct Submission   |            |                  |
|           | JOURNAL    | Submitted (18-FEB-1997) CLONTECH Laboratories, Inc., 1020 East Meadow Circle, Palo Alto, CA 94303-4230, USA   |            |                  |
| COMMENT   |            | This vector can be obtained from CLONTECH Laboratories, Inc., 1020 East Meadow Circle, Palo Alto, CA 94303-4230, USA. To place an order, call (415) 424-8222 or (800) 662-2566, extension 1.<br>International customers, please contact your local distributor or subsidiary. For technical information, call (415) 424-8222 or (800) 662-2566, extension 3.      |            |                  |
|           |            | This sequence has been compiled from information in the sequence databases, published literature and other sources, together with partial sequences obtained by CLONTECH. If you suspect there is an error in this sequence, please contact CLONTECH's Technical Support Department at (415) 424-8222 or (800) 662-2566, extension 3 or E-mail TECH@CLONTECH.COM. |            |                  |
| FEATURES  |            | Location/Qualifiers   |            |                  |

## FEATURES

Location/qualifiers

```
source      1. .6786
             /organism="unidentified cloning vector"
             /db_xref="taxon:45196"
gene         152. .811
             /gene="HIS3"
             /gene="HIS3"
             /codon_start=1
             /product="HIS3"
             /protein_id="AAB64382.1"
             /db_xref="GI:2190693"
             /translation="MTEOKALVYRITNETKIQIAISLKGGLPIAEHSIPEKEAEVA
             EQATQOQVINVHTIGLIDHMLAHKHSGLSIVLCIGDLHTIDHHTEDCGIAIGO
             AFKEALLARVKRRFGSPALDEALSRAYVDLSNRYAVAVELIOREKVDLSCEMP
             HFELEFAEASRIFLHYDCLRGKNDHHRSESAFALAVAIREATSPGNTDVPSTKVL
             M"
gene        2729. .3532
             /gene="URA3"
             /gene="URA3"
             /gene="URA3"
             /codon_start=1
             /product="URA3"
             /protein_id="AAB64383.1"
             /db_xref="GI:2190696"
             /translation="MSKATYKERAAHPSPFAKLFINHEKOTNLCAIDVETTEL
             LEIVELGPRICLKTTHVDILTFDSMEGTVPKPLKLSAKYNFLFEDRKPADIGNTVK
             IOYSAGVYRIAEADITNAHGVVPGIVSGIKOAEVEMKEPGLIMLAELISCKGSL
             TGEYTKGYDIAKSDKDFVIQPIAQRDMGSRDGYWMLINTPEVGLDDKDALGOQYR
             TYDDVVSSTSDIITVGRGLFAKGRDAKVEGERYRKAKWEAYLNRCSQON"
gene        4846. .5706
             /gene="AmpR"
             /complement(4846. .5706)
             /function="ampicillin resistance"
             /codon_start=1
             /product="beta-lactamase"
             /protein_id="AAB64384.1"
             /db_xref="GI:2190697"
             /translation="MSIOHFRVALIPFPFAFLPVAHPETLVKVADEADQLGARVY
             IELDLSGKILIESFRPEFRPMSTFKVLGAVISRIDAGQDQLGRRIHYXSNQDVE
             YSPYTKKHLTDGMTVRELCSAATITMSDNTNANLITITGPKRLTFLNMGDHVRL
             DRNEPLNEAIFPDERDITPAVMATTLKRLGELTLASRQQLIDMEADENVAPL
             LRSALPAGWFIADKSGAGERSKGIITAAIGLPPDKPSRIYIYTTGSAOTMDERNQIA
             EIGASLIIKHM"
BASE COUNT 1782 a 1590 c 1749 g 1665 t
ORIGIN
Query Match      35.6%; Score 2830.6; DB 56; Length 6786;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2836; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 5091 ccagctgcatatgaatcgcccaacgcgcgggagagagcggttcgcatattggcgctc 5150
      |||||||
DB 3847 CCAGCTGCATTATTAAGATTCGCCCAACGCGGAGAGCGGTTTGCGTATTTGGCCCTC 3906
QY 5151 ttccgcctctcgcctcactcactcgcgtcgctcgctcgctcgctcgctcgctcgctc 5210
      |||||||
DB 3907 TTCCGCTTCTCCTCCTACTACTCTACTCTGCGCTCGCTCGCTCGCTCGCGAGCGGATC 3966
QY 5211 agctcaactcaaaagcggttaactcgatlatccacagaaatcaggggagatacgcaggaagaa 5270
      |||||||
DB 3967 AGCTCACTCAAAAGCGGTTATACGGTTATCCACAGATCAGGGGATTAACGACGAAGAA 4026
QY 5271 catgtgagcaaaagcgagcaaaaagcgagaaacctgtaaaaagcgcgcttctgtgcgtc 5330
      |||||||
DB 4027 CATGTGAGCAAAAGCGAGCAAAAGCGCAGAACCGTAAAGAGCCCGCTTGTGGGCTT 4086
QY 5331 ttccatagagctcgcgcctcgaagagcatcacaaaaatcgagcgctcaagtcagaggtg 5390
      |||||||
DB 4087 TTTCGATAGGCTCCGCCCTGACGAGCATCACAAAAATCGAGCTCAAGTCAAGAGTG 4146
QY 5391 ggcgaacccgacaggaactaaagataccagcgcttcccccctggaagctccctcgtgcg 5450
      |||||||
```

```
DB 4147 GCGAAACCCGACAGACTATTAAGATACAGAGCGTTTCCCTGGAAAGTCCTCGTGCG 4206
QY 5451 ctccctcgttccgacctgcgccttaccggataaccttccgccttctcccttcgggaag 5510
      |||||||
DB 4207 CTCTCCGTGTCGACCCCTGCGCTTACCGGATTAACCTGTGCGCTTCTCCCTGGGAG 4266
QY 5511 cgtgcgcttctcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatc 5570
      |||||||
DB 4267 CGTGGCGCTTCTCATATAGCTACACGCTAGCTATCTCATGTTGGGTAGCTGTTCCGCTC 4326
QY 5571 caagctggcgctgtgtgcaagaaacccccgcgttcaagcccgacgcgctgcctatccggtaa 5630
      |||||||
DB 4327 CAACCTGGCGCTGTGACAGCAACCCCGCTTCAACCCGACCGCTGCGCTTATCCGGTAA 4386
QY 5631 ctatcgctctgagtcgaacccccgcgttcaagaaacgaactatcgccatctgcaagcgcacttg 5690
      |||||||
DB 4387 CTATCGTCTTGAGTCCAAACCCGCTTAAACACAGACTTATCCGACTGTGACAGCCACTCG 4446
QY 5691 taacagatlagcagagcgagatagttagcggtgtcacagagctcttgaagtggtgccc 5750
      |||||||
DB 4447 TAACAGGATTAGCAGACGAGAGATATTAGCGGCTGCTACAGAGTCTTGAAGTGTGTGCC 4506
QY 5751 taactacggtctacacatagaagacagatattggtatctgcgtctgtgaagccaagttac 5810
      |||||||
DB 4507 TAACCTACGCTACACTAGAAAGACAGATTTGATCTGCGCTCTGCTGAAGCCAGTTAC 4566
QY 5811 ctccggaaaaagagctgtgtgctctgtatcccggaacaaaccccgctggttagcggtg 5870
      |||||||
DB 4567 CTTCGGAAGAAAGAGTGTGTGATCGCTTATCGGCAACCAACCCGCTGTGTACCGGTGG 4626
QY 5871 ttttttgcttgcgaagcagcagatattacgcgcagaaaaaagagatctaaagaatccctt 5930
      |||||||
DB 4627 TTTTGTGTTTGTCAGACGACGAGATTACGCGCAGAAAAAGATCTTAAGAAATCTCTT 4686
QY 5931 gatcttctcagcgggtctgcagctcagctggaacgaacacacacacacacacacacacacac 5990
      |||||||
DB 4687 GATCTTTTCTACGGGGTCTGACGCTCAGTGGAAGAAACACAGTTAAGGATTTTGGT 4746
QY 5991 catgagattatcaaaaagagatcttcacatagatccctttaaataaaatgaaatttaa 6050
      |||||||
DB 4747 CATCAGATTATCAAAAAGAGATCTTACCTGATCCCTTTTAAATTAATAAATCAAACTTTTAA 4806
QY 6051 atcaatctaaagatatataagtaaaacttgctgcagcaatcgaatcgaatcgaatcgaatcga 6110
      |||||||
DB 4807 ATCATCTTAAGTATATATAGTAACTTGTGCTGACAGATTACCAATGCTTAAATCAAGTGA 4866
QY 6111 ggcacctatctcagcgatctgtctatctcgttcatcatcatcatcatcatcatcatcatcat 6170
      |||||||
DB 4867 GGCACCTATCTCAGCGCATCTGTCTATTTGCTTCATTCATAGTTGCTGAGACTCCCGTCT 4926
QY 6171 gtgataactcagatacaggaagggcttaccatctggccccaagtgctgcaatgataccgcg 6230
      |||||||
DB 4927 GTAGATTAACATACGATACGAGGAGGCTTACCATCTGCGCCCAAGTCTCAATGATACCGCG 4986
QY 6231 agacccacgcctcaccgcgctcagattatcagcaaaacacacacacacacacacacacacac 6290
      |||||||
DB 4987 AGACCCACGCTCACCGCTCAGATTTATACGACATTAACACGACGACCGGAAAGGCGCGA 5046
QY 6291 ggcgagaagtggtctcgaactatccatccatccatccatccatccatccatccatccatccat 6350
      |||||||
DB 5047 GCGCAGAAAGTGTCCTCAGCTTATCTCGGCTCCATCCATCCATCCATCCATCCATCCATCC 5106
QY 6351 agcttagagtaagtaagtgctgcgaagtaagtttgcgcaagctgtgtgccaattgctaagg 6410
      |||||||
DB 5107 AGCTTAGAGTAAGTAGTTCGCGAGTAATTAATTTGCGCAAGGTTTGCATTGCTTAAGG 5166
QY 6411 catcggtgtcagcgcgtcgttgtgtatggtcttaactaagctcgcggttcccaagatc 6470
      |||||||
DB 5167 CATCGTGTCACGCTCGTCTGTTGATGCTTACGCTCCGCTTCCCAACGATC 5226
QY 6471 aagcgagttacatgataccccaatgctgtgtgcaaaaaagcggttagctcctcgtctcc 6530
      |||||||
DB 5227 AAGCGCAGTTACATGATGCCCATGTGTGTCAAAAAAGCGGTTAAGTCTTCCGCTCTCC 5286
```

QY 6531 gatcgttgcagaagtaagtttggccgcagtgatcactcaatcgttattgcaagcactgca 6590  
|||||  
Db 5287 gatcgttgcagaagtaagtttggccgcagtgatcactcaatcgttattgcaagcactgca 5346  
6591 taattctctactcgtcactcgttaagatgcttctcgtgactggtgtaactcaac 6650  
|||||  
Db 5347 taaattctctactcgtcactcgttaagatgcttctcgtgactggtgtaactcaac 5406  
6651 caatctactcgtgagaatgattatgcccgcagagttgctcttgcggcggtcgaatacg 6710  
|||||  
Db 5407 caatctactcgtgagaatgattatgcccgcagagttgctcttgcggcggtcgaatacg 5466  
6711 ggaataatccgcgcacataagcaagacttaaaagtctcatalcattggaaaaagtcctc 6770  
|||||  
Db 5467 ggaataatccgcgcacataagcaagacttaaaagtctcatalcattggaaaaagtcctc 5526  
6771 gggcgcaaaactctcgaagatcctcagcgtgttgagatccagttcgatgtaaccactcg 6830  
5527 gggcgcaaaactctcgaagatcctcagcgtgttgagatccagttcgatgtaaccactcg 6830  
6831 tgcacccaactgactcctcagcactccttactcaccagcgttctgggtgagcaaaaac 6890  
|||||  
Db 5587 tgcacccaactgactcctcagcactccttactcaccagcgttctgggtgagcaaaaac 5646  
6891 aggaagcgcaaaatgcgcgcaaaaaggaataagggcgacaagaaatgltgaatactcat 6950  
5647 aggaagcgcaaaatgcgcgcaaaaaggaataagggcgacaagaaatgltgaatactcat 5706  
6951 actcttcttctcaataatattgaagcattatcagagttatgtcctcagaagcgata 7010  
7007 actcttcttctcaataatattgaagcattatcagagttatgtcctcagaagcgata 7010  
7011 cataattgaaatgatttgaaaaaataaacaataaggggttcggcgacattcccgaa 7070  
5767 catattgaaatgatttgaaaaaataaacaataaggggttcggcgacattcccgaa 5826  
7071 agtgcacactgacgtcctaagaacacattatatacctgaataacctaataaataagcg 7130  
5827 agtgcacactgacgtcctaagaacacattatatacctgaataacctaataaataagcg 5886  
7131 tatacgaagggccttctcgtcgcggttctcgtgataagcggtgaagaacctctgacacat 7190  
5887 tatcaccagggccttctcgtcgcggttctcgtgataagcggtgaagaacctctgacacat 5946  
7191 gcaacgtcccgagagacgtgcacacagctgtctgaacgagatgcgggagacagcaagcccg 7250  
5947 gcaacgtcccgagagacgtgcacacagctgtctgaacgagatgcgggagacagcaagcccg 6006  
7251 tcaaggcgagtcagcgaggtgttgcgcggtgtcgcgggtctgacttaactatgcgcgacaga 7310  
6007 tcaaggcgagtcagcgaggtgttgcgcggtgtcgcgggtctgacttaactatgcgcgacaga 6066  
7311 gcaagatgactgagagatgcacacataatgcgtgtgaaatacgcgcacaatgtagtaagag 7370  
6067 gcaagatgactgagagatgcacacataatgcgtgtgaaatacgcgcacaatgtagtaagag 6126  
7371 aaaaatacgcacatacagaatgtaaaacttaataatctgacttaactatgcgcgacaga 7430  
6127 aaaaatacgcacatacagaatgtaaaacttaataatctgacttaactatgcgcgacaga 6186  
7431 tctttaaatacagtcacttcttaacaaatagcgcaaatctcgcaaaaatccctataactca 7490  
6187 tctttaaatacagtcacttcttaacaaatagcgcaaatctcgcaaaaatccctataactca 6246  
7491 aaagaatacagtcacttcttaacaaatagcgcaaatctcgcaaaaatccctataactca 7550  
6247 aaagaatacagtcacttcttaacaaatagcgcaaatctcgcaaaaatccctataactca 6306  
7551 aaagaatacagtcacttcttaacaaatagcgcaaatctcgcaaaaatccctataactca 7610  
6307 aaagaatacagtcacttcttaacaaatagcgcaaatctcgcaaaaatccctataactca 6366

QY 7611 cgtgaaccatcacccatcaaatgatttctgggtgcaggtgcgtgaagaacactaatcgg 7670  
|||||  
Db 6367 cgtgaaccatcacccatcaaatgatttctgggtgcaggtgcgtgaagaacactaatcgg 6426  
7671 aacctaaaggaagcccccagatttagaacttgaacggtgaagagccggcgcaacgttgcgaga 7730  
6427 aacctaaaggaagcccccagatttagaacttgaacggtgaagagccggcgcaacgttgcgaga 6486  
7731 aaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaagga 7790  
6487 aaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaagga 6546  
7791 ctgcggttaac 7850  
6547 ctgcggttaac 6606  
7851 tgcgcatcaggtcagcgaacgttgggaaggaaggaaggaaggaaggaaggaaggaaggaagga 7910  
6607 tgcgcatcaggtcagcgaacgttgggaaggaaggaaggaaggaaggaaggaaggaaggaagga 6666  
7911 cgcagctgcgcaggg 7935  
6667 cgcagctgcgcaggg 6691

RESULT 11  
CVGEM13ZP  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS

SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

REMARK  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

REMARK  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

COMMENT

FEATURES  
source  
misc\_feature  
misc\_feature  
promoter  
misc\_feature

1. 3179  
/organism="Cloning vector pGEM-13zf(+)"  
/db\_xref="taxon:90121"  
1. 43  
/note="T7 transcription initiation site"  
49..68  
/note="Sp6 promoter"  
51

Cloning vector pGEM-13zf(+).  
Cloning vector pGEM-13zf(+).  
Official sequence; vectors.  
1 (bases 1 to 3179)  
Technical Services.  
Direct Submission  
Submitted (23-MAR-1992) Technical Services, Promega Corporation,  
2800 Woods Hollow Road, Madison, WI 53711-5399, USA  
revised by [2]  
2 (bases 1 to 3179)  
Technical Services.  
Direct Submission  
Submitted (28-MAY-1993) Technical Services, Promega Corporation,  
2800 Woods Hollow Road, Madison, WI 53711-5399, USA  
revised by [3]  
3 (bases 1 to 3179)  
Kendrick, M.  
Direct Submission  
Submitted (02-JUL-1999) Kendrick M., Technical Writing, Promega  
Corporation, 2800 Woods Hollow Road, Madison, WI 53711, USA  
On Jul 8, 1999 this sequence version replaced gi:58170.  
See X65300-X65335 for related vector sequences  
This vector can be obtained from Promega Corporation, Madison, WI.  
Call one of the following numbers for order or technical  
information:  
Order or Technical 800-356-9526  
In Wisconsin 800-356-9526  
Outside U.S. 608-274-4330.  
Location/Qualifiers  
1. 3179  
/organism="Cloning vector pGEM-13zf(+)"  
/db\_xref="taxon:90121"  
1. 43  
/note="T7 transcription initiation site"  
49..68  
/note="Sp6 promoter"  
51



|    |      |          |          |         |         |        |        |          |          |        |        |         |         |      |        |     |      |     |      |     |         |      |
|----|------|----------|----------|---------|---------|--------|--------|----------|----------|--------|--------|---------|---------|------|--------|-----|------|-----|------|-----|---------|------|
| QY | 6651 | caagtcac | ttc      | tgggaat | ta      | gt     | gt     | tgt      | ggg      | gcga   | ccga   | gt      | gt      | cct  | gt     | ccg | gcg  | tca | at   | ta  | gc      | 6710 |
|    |      |          |          |         |         |        |        |          |          |        |        |         |         |      |        |     |      |     |      |     |         |      |
| Db | 1808 | CAAGTCAT | TTTGAAAT | AGTGTAT | GGCGAC  | CGAGTT | GGCTTG | CCCGGCG  | GTAAATAC | GC     |        |         |         |      |        |     |      |     |      |     | 1867    |      |
| QY | 6711 | gga      | ta       | at      | cc      | gc     | gc     | cc       | at       | ta     | gc     | ag      | aa      | ct   | ta     | aa  | gc   | t   | ca   | tc  | a       | 6770 |
|    |      |          |          |         |         |        |        |          |          |        |        |         |         |      |        |     |      |     |      |     |         |      |
| Db | 1868 | GGATTAAT | ACCGCGC  | ACATAT  | AGCA    | GAAC   | TTTAA  | AGTCTCAT | TTAT     | TGGAA  | AAAC   | GTCTTC  |         |      |        |     |      |     |      |     | 1927    |      |
| QY | 6771 | gggg     | gc       | aa      | aa      | aa     | ct     | tc       | aa       | gc     | t      | tc      | aa      | gc   | gt     | gt  | ta   | gc  | at   | tc  | gc      | 6830 |
|    |      |          |          |         |         |        |        |          |          |        |        |         |         |      |        |     |      |     |      |     |         |      |
| Db | 1928 | GGGGC    | CAAAA    | CTCTCA  | AGAT    | CTTAC  | CGCTTG | TGAGAT   | CCAGTT   | CGAT   | TGTA   | AA      | CCCA    | CTTC |        |     |      |     |      |     | 1987    |      |
| QY | 6831 | tgc      | ac       | cc      | aa      | ct     | gc     | a        | ct       | tc     | ca     | gc      | a       | ct   | tc     | tc  | ta   | cc  | aa   | gc  | gc      | 6890 |
|    |      |          |          |         |         |        |        |          |          |        |        |         |         |      |        |     |      |     |      |     |         |      |
| Db | 1988 | TGCACCA  | CA       | NCAT    | GTAT    | TCAC   | AT     | CTTTACT  | TTTCCAC  | GCGTTT | CTGGTG | TAGCA   | AAAA    | AC   |        |     |      |     |      |     | 2047    |      |
| QY | 6891 | agga     | ag       | gc      | aa      | aa     | at     | gc       | gc       | aa     | aa     | aa      | gg      | gc   | gc     | gc  | gc   | gc  | gc   | gc  | gc      | 6950 |
|    |      |          |          |         |         |        |        |          |          |        |        |         |         |      |        |     |      |     |      |     |         |      |
| Db | 2048 | AGGA     | AGGCA    | AAAA    | CCGCGCA | AAAA   | GGGA   | AA       | TAGGCG   | CA     | CGCA   | CGGA    | AA      | GT   | GAAT   | CA  | TCAT |     |      |     | 2107    |      |
| QY | 6951 | ac       | ct       | tc      | cc      | tt     | tc     | aa       | ta       | ta     | ta     | gc      | gc      | at   | tc     | tc  | aa   | gc  | gc   | gc  | gc      | 7010 |
|    |      |          |          |         |         |        |        |          |          |        |        |         |         |      |        |     |      |     |      |     |         |      |
| Db | 2108 | ACTCTT   | CTCTTT   | CAAT    | AT      | TAT    | TGA    | MA       | CA       | TTAT   | CA     | GGGTTAT | GTCTCAT | GTAG | CGGATA |     |      |     |      |     | 2167    |      |
| QY | 7011 | ca       | ta       | tt      | gc      | aa     | ta     | ta       | aa       | aa     | aa     | aa      | aa      | aa   | aa     | aa  | aa   | aa  | aa   | aa  | aa      | 7070 |
|    |      |          |          |         |         |        |        |          |          |        |        |         |         |      |        |     |      |     |      |     |         |      |
| Db | 2168 | CATAT    | T        | TGAAT   | GTAT    | T      | TAG    | AAAA     | ATAA     | CA     | AA     | TAT     | AGGG    | TTCC | GGCA   | CA  | AT   | T   | T    | T   | CCCGAAA | 2227 |
| QY | 7071 | agt      | gc       | ca      | cc      | tg     | aa     | gc       | gt       | ct     | aa     | ga      | aa      | cc   | ta     | ta  | ca   | gc  | aa   | ta  | aa      | 7130 |
|    |      |          |          |         |         |        |        |          |          |        |        |         |         |      |        |     |      |     |      |     |         |      |
| Db | 2228 | AGG      | CCAC     | CTG     | AC      | GTCT   | TAA    | AAAA     | CCAT     | TAT    | TAT    | CA      | TAT     | CA   | CTAT   | TAA | CCAT | TAT | AAAA | TAT | AGCG    | 2287 |
| QY | 7131 | ta       | ta       | ca      | gc      | ag     | gc     | cc       | ct       | tc     | gc     | gc      | gt      | tc   | gc     | gc  | gt   | tc  | gc   | gc  | gc      | 7190 |
|    |      |          |          |         |         |        |        |          |          |        |        |         |         |      |        |     |      |     |      |     |         |      |
| Db | 2288 | TAT      | CA       | GC      | AG      | GC     | CC     | CTT      | CG       | T      | CG     | T       | CG      | T    | CG     | T   | CG   | T   | CG   | T   | CG      | 2347 |
| QY | 7191 | gc       | ag       | ct      | cc      | gc     | gc     | gc       | gc       | gc     | gc     | gc      | gc      | gc   | gc     | gc  | gc   | gc  | gc   | gc  | gc      | 7250 |
|    |      |          |          |         |         |        |        |          |          |        |        |         |         |      |        |     |      |     |      |     |         |      |
| Db | 2348 | GC       | AG       | CT      | CC      | CG     | GA     | AC       | CG       | GT     | CA     | CA      | CTT     | CT   | GT     | TAA | GC   | CG  | GA   | AC  | CA      | 2407 |
| QY | 7251 | tc       | ag       | gc      | gc      | gc     | gc     | gc       | gc       | gc     | gc     | gc      | gc      | gc   | gc     | gc  | gc   | gc  | gc   | gc  | gc      | 7310 |
|    |      |          |          |         |         |        |        |          |          |        |        |         |         |      |        |     |      |     |      |     |         |      |
| Db | 2408 | TC       | AG       | GC      | GC      | GT     | CA     | GC       | GG       | GT     | T      | TG      | CG      | GG   | GT     | T   | TG   | CG  | GG   | GT  | T       | 2467 |
| QY | 7311 | gc       | ag       | at      | tg      | ta     | ct     | gc       | gc       | gc     | gc     | gc      | gc      | gc   | gc     | gc  | gc   | gc  | gc   | gc  | gc      | 7370 |
|    |      |          |          |         |         |        |        |          |          |        |        |         |         |      |        |     |      |     |      |     |         |      |
| Db | 2468 | GC       | AG       | AT      | T       | TGA    | CT     | G        | AT       | G      | AT     | G       | AT      | G    | AT     | G   | AT   | G   | AT   | G   | AT      | 2527 |
| QY | 7371 | aaa      | at       | cc      | gc      | at     | ta     | gc       | aa       | at     | tg     | ta      | aa      | ct   | ta     | at  | tt   | tg  | ta   | aa  | at      | 7430 |
|    |      |          |          |         |         |        |        |          |          |        |        |         |         |      |        |     |      |     |      |     |         |      |
| Db | 2528 | AAAA     | T        | CCG     | AT      | CA     | TGA    | AA       | AT       | T      | TG     | AA      | CG      | TT   | T      | TG  | AA   | AA  | T    | T   | T       | 2587 |
| QY | 7431 | tg       | ta       | at      | ca      | gc     | ta     | ct       | tt       | tc     | aa     | cc      | aa      | ta   | gc     | gc  | aa   | aa  | ct   | cc  | ta      | 7490 |
|    |      |          |          |         |         |        |        |          |          |        |        |         |         |      |        |     |      |     |      |     |         |      |
| Db | 2588 | TG       | T        | TAA     | T       | CAG    | CT     | CA       | T        | T      | T      | T       | T       | T    | T      | T   | T    | T   | T    | T   | T       | 2647 |
| QY | 7491 | aaa      | ga       | ta      | gc      | gc     | gc     | gc       | gc       | gc     | gc     | gc      | gc      | gc   | gc     | gc  | gc   | gc  | gc   | gc  | gc      | 7550 |
|    |      |          |          |         |         |        |        |          |          |        |        |         |         |      |        |     |      |     |      |     |         |      |
| Db | 2648 | AA       | GA       | AT      | TAA     | CCG    | AG     | TAA      | GG       | T      | AG     | T       | GG      | T    | T      | T   | T    | T   | T    | T   | T       | 2707 |
| QY | 7551 | aag      | aa       | gc      | tg      | aa     | ct     | cc       | aa       | gc     | tc     | aa      | gc      | gc   | gc     | gc  | gc   | gc  | gc   | gc  | gc      | 7610 |
|    |      |          |          |         |         |        |        |          |          |        |        |         |         |      |        |     |      |     |      |     |         |      |
| Db | 2708 | AA       | GA       | AC      | T       | G      | AC     | T        | G        | AC     | T      | G       | AC      | T    | G      | AC  | T    | G   | AC   | T   | G       | 2767 |
| QY | 7611 | cg       | ta       | gc      | cc      | ta     | cc     | ta       | ta       | gc     | tt     | tc      | tg      | gc   | gc     | gc  | gc   | gc  | gc   | gc  | gc      | 7670 |
|    |      |          |          |         |         |        |        |          |          |        |        |         |         |      |        |     |      |     |      |     |         |      |
| Db | 2768 | CG       | T        | GA      | AC      | CA     | T      | AC       | CC       | T      | T      | AA      | T       | CA   | AT     | T   | T    | T   | T    | T   | T       | 2827 |
| QY | 7671 | aa       | cc       | cc      | ta      | aa     | gg     | gc       | gc       | gc     | gc     | gc      | gc      | gc   | gc     | gc  | gc   | gc  | gc   | gc  | gc      | 7730 |
|    |      |          |          |         |         |        |        |          |          |        |        |         |         |      |        |     |      |     |      |     |         |      |
| Db | 2828 | AA       | CC       | CC      | TAA     | AG     | G      | G        | A        | G      | G      | A       | G       | G    | A      | G   | G    | A   | G    | G   | A       | 2887 |

|            |   |   |      |
|------------|---|---|------|
| Qy         | 7731  | aaggaagggaagaagcgaaagagagcgagcgctcgaagtcgtgcaagtgtaacggtcaacg   | 7790 |
| Db         | 2888  | AAGGAGGAGGAGAGAAAGCGAAAGGAGCGCGGCGCTTAGGGCGTGGCAAGTACGGGTACAG   | 2947 |
| Qy         | 7791  | ctgcgcgtgaaccaccacaccccgccgcgtctaaatgcgcgcgtctaaagagcgctgcgcgat | 7850 |
| Db         | 2948  | CTGGCGCGTAACACACACACCGCGCGCTTAATATCGCGCGCTACAGGCGCGGT----CGAT   | 3003 |
| Qy         | 7851  | tgcgcattcaagctcagcgaactgcttggaagagcgatcgctgcggcccttcgcataata    | 7910 |
| Db         | 3004  | TGCGCATTCAGGCTGCGCAACTGTGGGAAAGGCGGATGGGTGGGCGCTCTTCGCTATTAA    | 3063 |
| Qy         | 7911  | cgcacagctgcgtcgaaggagggagggg                                    | 7935 |
| Db         | 3064  | CGCCACGCTGGCGGAAGGGGGAATGTG                                     | 3088 |
| RESULT     | 12  |   |      |
| CVPGEM32P  |   |   |      |
| LOCUS      | CVPGEM32P   | 3197 bp   | DNA  |
| DEFINITION | Cloning vector pGEM-3zf(+).   |   |      |
| ACCESSION  | X65306  |   |      |
| VERSION    | X65306.2  | GI:5531232  |      |
| KEYWORDS   | cloning vector; multiple cloning site; phage fl region; promoter.   |   |      |
| SOURCE     | Cloning vector pGEM-3zf(+).   |   |      |
| ORGANISM   | Cloning vector pGEM-3zf(+)  |   |      |
| REFERENCE  | 1 (bases 1 to 3197)   |   |      |
| AUTHORS    | Solomon,L.R., Massom,L.R. and Jarrett,H.W.  |   |      |
| TITLE      | Enzymatic syntheses of DNA-silicas using DNA polymerase   |   |      |
| JOURNAL    | Anal. Biochem. 203 (1), 58-69 (1992)  |   |      |
| MEDLINE    | 92398067  |   |      |
| REFERENCE  | 2 (bases 1 to 3197)   |   |      |
| AUTHORS    | Technical,Services.   |   |      |
| TITLE      | Direct Submission   |   |      |
| JOURNAL    | Submitted (23-MAR-1992) Technical Services, Promega Corporation,  |   |      |
|            | 2800 Woods Hollow Road, Madison, WI 53711-3599, USA   |   |      |
| REMARK     | revised by [3]  |   |      |
| REFERENCE  | 3 (bases 1 to 3197)   |   |      |
| AUTHORS    | Technical,Services.   |   |      |
| TITLE      | Direct Submission   |   |      |
| JOURNAL    | Submitted (28-MAY-1993) Technical Services, Promega Corporation,  |   |      |
|            | 2800 Woods Hollow Road, Madison, WI 53711-5399, USA   |   |      |
| REMARK     | revised by [4]  |   |      |
| REFERENCE  | 4 (bases 1 to 3197)   |   |      |
| AUTHORS    | Kenefick,K.   |   |      |
| TITLE      | Direct Submission   |   |      |
| JOURNAL    | Submitted (14-JUL-1999) Kenefick K, Technical Services, Promega Corporation,  |   |      |
|            | 2800 Woods Hollow Road, Madison, WI 53711-5399, USA   |   |      |
| COMMENT    | On Jul 19, 1999 this sequence version replaced gi:58203.<br>See X65300-X65335 for related vector sequences<br>This vector can be obtained from Promega Corporation, Madison, WI<br>Call one of the following numbers for order or technical<br>information:<br>Order or Technical 800-356-9526<br>In Wisconsin 800-356-9526<br>Outside U.S. 608-274-4330. |   |      |
| FEATURES   |   |   |      |
| source     | Location/Qualifiers   |   |      |
|            | 1..3197   |   |      |
|            | /organism="Cloning vector pGEM-3zf(+)"  |   |      |
|            | /db_xref="taxon:90112"  |   |      |
|            | 1   |   |      |
|            | /note="T7 transcription initiation site"  |   |      |
|            | 5..61   |   |      |
|            | /note="multiple cloning sites"  |   |      |
|            | 67..86  |   |      |
|            | /note="SP6 promoter"  |   |      |
|            | 69  |   |      |
|            | /note="SP6 transcription initiation site"   |   |      |
|            | 94..323   |   |      |
|            | /note="lac operon sequence"   |   |      |
|            | complement(106..108)  |   |      |
|            | /gene="lacZ"  |   |      |

```

gene      /note="lac2_start_codon"
           complement(106. ,108)
gene      /gene="lac2"
           complement(1265. ,2125)
gene      /gene="bla"
           complement(1265. ,2125)
CDS       /gene="bla"
           /codon_start=1
           /transl_table=1
           /product="beta-lactamase"
           /protein_id="CAA6398.2"
           /db_xref="gi:553123"
           /translation="MSIOHERVALIPFEAPCLVPFAHPERTLVKKAEDOLGARVGE
           IELDNGSKILIESRPREPRPMSTFETVLLCGAYLSRIDAQEOELGRRIIHQNDLVG
           YSPTEKRLHIDGTAVRELCSAITTSDDTANLLITIGCKELTALFHHNDVHTRI
           DRWPEELNEAIPIERDPTMPVAMATLRLKLLTSELTLFASROOLIDMENDKDAAGP
           LRSRLPAGWFIADKRSVGSGERGSRGIIAALGPDGPRSLIVITYGSOATMDERNROILA
           EIGASLIKHW"
misc_feature 2562. .3017
           /note="phage fl region"
misc_feature 3018. ,3178
           /note="lac operon sequence"
promoter    join(3181. ,3197,5)
           /note="T7 promoter"
BASE COUNT      822 a      785 c      817 g      773 t
ORIGIN

```

|                            |        |                |           |              |
|----------------------------|--------|----------------|-----------|--------------|
| Query Match                | 35.4%; | Score 2809.4;  | DB 56;    | Length 3197; |
| Best Local Similarity      | 99.5%; | Pred. No. 0;   |           |              |
| Matches 2830; Conservative | 0;     | Mismatches 11; | Indels 4; | Gaps 1;      |

|    |      |   |      |
|----|------|---|------|
| OY | 5091 | ccagcgcgcatatagatgaatcggcccaacgcgcgcgggagagcggttgcgcatlbggcgc       | 5150 |
| Db | 266  | CCACCTCCTATTATGATCGGCCCAAGCGCGGGAGAGCGGCTTTGCGTATTGGAGCCTC          | 325  |
| OY | 5151 | ttccgcgttcctgcgtcaactgaactcgcgtctgcgcgcgtgcgttcgtgcgcgcgcgcgtatc    | 5210 |
| Db | 326  | TTCCGGCTTCTCGCTACCTACTACTCGCTGGCGTCCGTCGTTCCGGCTGGCGGCGAGCGGTATC    | 385  |
| OY | 5211 | agctcaactcaaaagcgcgttaatacgcgttatccacgaatatcaaggataaagcaaganaa      | 5270 |
| Db | 386  | AGCTCACCTCAAAAGCGGGTATATACGGTTATCCACAGATCAGGGAGTAAACGACGAAAGAA      | 445  |
| OY | 5271 | catgtgagcaaaagcgccagcaaaagcccgagaaacggttaaagccgcgcgttgcgtgcgt       | 5330 |
| Db | 446  | CATGTGAGCAAAAGCCACGCAAAAAGGCCACGGAACCGTAAAAAGCCGGCTGTGCGGCTT        | 505  |
| OY | 5331 | tttccatagagctccgcgcgccttcgacgaagatcaacaataatgaagctcaagtcgaagtg      | 5390 |
| Db | 506  | TTTCCATAGGCTCCGCCCCCTGACGACGATCACAAAATCGACCTCAAGTCAGAGGTG           | 565  |
| OY | 5391 | gcgaacccgcagcagactataagataaccagcgcgtttcccccctgaaagctccctcgtgcg      | 5450 |
| Db | 566  | GCGAACCCGACAGGACTATAAAGATACCAAGCGGTTCCCCCTCGAAGACTCCCTCGTGC         | 625  |
| OY | 5451 | ctcccccgttcgaacctgcgcgttcacgaatacctgcgcgttcccttcctcccttcggag        | 5510 |
| Db | 626  | CTTCTCGTGTTCGACCTGCGCCGTACCGGATACCTGTCGCGCTTTCCTTCCTTGCGGAG         | 685  |
| OY | 5511 | cgtagcgttcttcaatgctcaacgcgcgttagatgatactcaagtlcgcgtgtagtgcgttcgc    | 5570 |
| Db | 686  | CGTGGCGCTTTCATATAGTCACGCTGTAGATATCAAGTTCCGGTGTAGGTCTTCCGTC          | 745  |
| OY | 5571 | caagcgtgcgtgtgtgcagaacacccccggttcaagccgaacgcgcgtgcgcctatccgtaa      | 5630 |
| Db | 746  | CAACCTGGCGTGTGTGACCAACCCCCCGTTACGCGCCGACCGCTCCGCTTATCCGGTAA         | 805  |
| OY | 5631 | ctatgcgccttgaagccaacccggttaagaagaactatccgaactgcgcgcgcgcgcgcgcgc     | 5690 |
| Db | 806  | CTATGCTCTTGAGTCCAAACCCGGTAAAGACAACGATTTATCGCCACTGGAGGACGCACTGG      | 865  |
| OY | 5691 | taaacagattagcagaagcgcgcgcgtatbtlaagcgcgtgcatacagaagtlctctgaagtggtgc | 5750 |

|    |      |   |      |
|----|------|---|------|
| Db | 866  | TAACAGGATTAGCAGACCGAGGATGATGAGCGGTGCTGTAAGAGTTCTTGAAGTGGTGGCC   | 925  |
| Qy | 5751 | taacaaagcctaacactaaagaagcaagatatttggtaactcgtctcgtcaagccagttac   | 5810 |
| Db | 926  | TAACTACGGCTACACTGAAAGAACAGATTTGGTATCTCGCTTGCTGTAAGCCAGTTAC      | 985  |
| Qy | 5811 | cttcgaaaaaagaatttggtagctcttgaatccgcgcaacaacaccgcgttggtagcgttgg  | 5870 |
| Db | 986  | CTTGGAAAAAAGAGTTGGTAGCTCTTGATCTCGGCAACCAACACCGCTGGAGGGTGG       | 1045 |
| Qy | 5871 | tttttttggtttgcgaagcagatatacgcgcagaaaaaaggatctcaagaatccctt       | 5920 |
| Db | 1046 | TTTTTTTGTGTCAGACGAGAGATATCGCGCCAAAAAAAAGATGTCAAGAAAGATCTTT      | 1105 |
| Qy | 5921 | gactcttctcagcgggtctbaagcctbaagcaggaagaaactaacgttaaggaatttgg     | 5990 |
| Db | 1106 | GATCTTTTCTACGGGGTCTGACGCTCAGTGGACGAAACTACGTTAAGGGATTTTGGT       | 1155 |
| Qy | 5991 | catgagatatacaaaaagaattcttcacctaagataccctttaaattaaatgaagttttaa   | 6050 |
| Db | 1166 | CATGAGATTATCAAAAAGATCTTTCAACCTTGATCTTTTAAATTAATTAATGAAGTTTAA    | 1225 |
| Qy | 6051 | atcaatctaaagatatataagtaaaacttggctcagcttaccaaagtctaactaagtga     | 6110 |
| Db | 1226 | ATCATCTTAAAGGTATATATAGTAAACTTGCTGTGACACTTACCAATGCTTAATACAGA     | 1285 |
| Qy | 6111 | ggcaactcttcacagcgtctgctcatttctgttaaccataagtttgcttaaccocccgt     | 6170 |
| Db | 1286 | GGCACCTATCTCACAGCTATGtCTATTTCGTTATCCATAGTTGCTGTGATCCCGCTGT      | 1345 |
| Qy | 6171 | gtagataactacagaataagggaggcttaccatcttgccccagtgctgcaatgatacccg    | 6230 |
| Db | 1346 | GTAGATTACTACGATACGGGGGCTTACATCTGGCCCCCAGTGGCAATGATATCCGCG       | 1405 |
| Qy | 6231 | agaaccacgctcaccggtctccagattatcagcaataaaccaagccagccggaaggccga    | 6290 |
| Db | 1406 | AGACCACAGCTCACCGGCTCCAGATTTTATCAGCAATAAACAGCCAGCGGAGGGCCGA      | 1465 |
| Qy | 6291 | ggcgaaggtggctccbaaacttataccgcttcacatcagttatataatttggtcggga      | 6350 |
| Db | 1466 | GGCGAAGATGGTCCCTCAACTTATTCGCGCTCCATCCAGTCTATTAAATGTGTGGCGGA     | 1525 |
| Qy | 6351 | agctagaataagtagtctcgcgcagttaaatatttgcgcaagcttggccaatgtctacag    | 6410 |
| Db | 1526 | AGCTAGAACTAGTAGTTCCGCACAGTTAATATTGGCCCAACGTTGTCATTTGCTACACG     | 1585 |
| Qy | 6411 | catcgttggtcacgcctcgtctgttggtagtgcatactcaagctcccgltcccaagatc     | 6470 |
| Db | 1586 | CATCGTGCTGCACGCTCGTCGTTGGTATGCTTATTCACGCTCCGGTCCCAACGATC        | 1645 |
| Qy | 6471 | aagcgcagttacatgatacccccaatgtttgtaaaaaaagcgttaagctccttcggtcccc   | 6530 |
| Db | 1646 | AAGGGATTTACATGATCCCCCATAGTTGTGCAAAAAAGGGTTAGCTCTTCCGTGGCTCC     | 1705 |
| Qy | 6531 | gactgttgcgaagtaagtttggccgcagtggtataactacatcagtgtagtcagactga     | 6590 |
| Db | 1706 | GATCGTTCTCGAAGTAGTTGGCCGAGTGTTATACATCAAGTTATGSGACGACTGCA        | 1765 |
| Qy | 6591 | taattctcttaactgtcatgtccatccgtaagaatgtcttctgtgaactgttgagttactaac | 6650 |
| Db | 1766 | TAAATCTCTTACTGCTCATAGCCATCCGTAAGATGCTTTCTGTGACTGGGTGAGTACTCACC  | 1825 |
| Qy | 6651 | caagtcattctggaataatagtgtagtggcgagccagattgctcttcgccggcgtaataag   | 6710 |
| Db | 1826 | CAAGTCATTCTGAAATAATAGTATAGGGGCCACCGAATGTTCTTTCCCGCGCTCAATACG    | 1885 |
| Qy | 6711 | ggaataaacccgcacatatagcagaactttaaaagtgctcatalcttggaaaaagttcttc   | 6770 |
| Db | 1886 | GGATATATACCCCGCCACATAGCAAGACTTTAAAAAGTGCATCATTTGGAAAAACGTTCTTC  | 1945 |
| Qy | 6771 | gggggaaaaactctcaagatcttaccgcgttggatccagttcgaatgaaccacatcg       | 6830 |



Db 1946 GGGGCGAAAACCTCTCAAGATCTTACCGCTGTTGAGATCCAGTTGAGTAACCCACTCG 2005  
 QY 6831 tgcacccaacatgctcagacatcttactctcaccagcgcttcctggctgagcaaaac 6890  
 Db 2006 TGCAACCCAACTGATCTGAGCATCTTTACTTTCATCTCAACAGCGTTCTCGGTGAGCAAAAC 2065  
 QY 6891 aggaagcagcaaaatgcgcgcaaaaaggaataagggcgacacggaatgltgaatacctac 6950  
 Db 2066 AGGAAGGCAAAATGCCGCAAAAAGGGAATAGGGCGACACCGAAATGTTGAATACTCAT 2125  
 QY 6951 acctctccttctcaatcttgaagcatatcagaagtcttctcaagaagcagata 7010  
 Db 2126 ACTCTCTCTTTTCAATTTATTTAGACATTTATCAGGGTTATTTGTTCTCATAGCGGATA 2185  
 QY 7011 cataattgaatglatltagaaaaataacaataggggtlccgcgcacatltcccgaaa 7070  
 Db 2186 CATATTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCGGCGACATTTCCCGGAAA 2245  
 QY 7071 agtgccacatgagctctcaagaacataataatcatatgacataacctataaataagcg 7130  
 Db 2246 AGTCCACCTACGCTCTAAGAAACCATTTATCATGACATTAACCTATAAAATAAGCGCG 2305  
 QY 7131 tatcacgaagcccttcgtccgcgcgttcggtgataaggtlgaaaacctctgacat 7190  
 Db 2306 TATACAGAGGCCCTTTCGCTCGCGGTTTGGTATGACGGTGAACCTCTGACACAT 2365  
 QY 7191 gcacgtcccgagagacggtcacagctgtctglaagcgagatgcgggagcgagacaagccg 7250  
 Db 2366 GCAGCTCCCGGAGAGAGGTGACAGCTTGCTGTAAACGGATGCGGGGAGAGCAAGCCCG 2425  
 QY 7251 tcagggcgagtcagcgaggtgtgtgcgggtgtcggggtcggttaactabgcgacata 7310  
 Db 2426 TCAGGGCGCGCTCAGCGGGGTGTGGGGGTGTGGGGCTGCTTAATGCGCGCATCAGA 2485  
 QY 7311 gcaagtgtactgagagagcgccacatatcggtgtgtaaatccgcgacagatggttaagag 7370  
 Db 2486 GCACATGTGACTGAGAGTGACACATATCGGGTGTAAATACCGCACAGATGGGTAAAGG 2545  
 QY 7371 aaaaacacgcatacgaagaattgtlaaacgtltaacttctgtaaaactcgcgtlaaatlt 7430  
 Db 2546 AAATATACCGCATCAGGAATTTGTAGCGTTATATTTTGTAAATTCGCCGTTAAATTTT 2605  
 QY 7431 tgttaaatcagctcaatttttaaccaatagcgcaaatcgcgcaaaatccctataatca 7490  
 Db 2606 TGTAAATACACTCATTTTAAACCAATAGGCCGAATCGGCAAAATCCCTTAATAATCA 2665  
 QY 7491 aaagaatagacgagataggttgaggtgtgttccagtttgaagaagagtcactatla 7550  
 Db 2666 AAGGAATAGACCGAGATAGGGTTGAGTGTGTTCAGTTTGGACACAGAGTCCACTATTA 2725  
 QY 7551 aagaacgtlgaatccaaatgcgaagcgcaaaaacgctltaacagggcagatgcccacta 7610  
 Db 2726 AAGAACTGAGCTCAACAGTCAAAAGGGGGAATAACCGTCTATCAGGGGATGGCCACGA 2785  
 QY 7611 cgtgaacatcaaccttaactaagtlttttgggttcgaggtgcgtaagactaactcgg 7670  
 Db 2786 CGTAAACATCACTCCCTATCAAGTTTGTGGGTGAGAGTGCCTGAAGACACTTAATCGG 2845  
 QY 7671 aaacctaaaggaagcccccgaattagaagcttgaacgaggaaggaagcggaagctgcaga 7730  
 Db 2846 AACCTTAAGGAGCCCCCGATTAGACTTGAACGCGGGAAGCCGCGCAACCTGGCGAGA 2905  
 QY 7731 aaggaaggaagaaagcgaaggaagcgcgctgaagggcgttcgaagtgaagcgttcacg 7790  
 Db 2906 AAGGAAGGAAGAAAGCAAGAGCGCGCGCTAGAGCGCTGCGCAAGTGAAGCGGTCAAG 2965  
 QY 7791 ctgcggtlaaccacacacccgcgcgttaatgcgcgcgtacagggcggtlccgcgcac 7850  
 Db 2966 CTGCGCTTAACACACACACCGCGCGCTTATGCGCGCTAACAGGGCGCGT---CCAT 3021  
 QY 7851 tcgacatcaggtcagcaaatgtttgggaagggcagatcggtgcggagcctcttcgctata 7910  
 Db 3022 TCGCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGTGGCGGCTCTTTCGCTATTA 3081

QY 7911 cgcacgtgctcagggggggggg 7935  
 Db 3082 CCCAGAGCTGGCAAGGGGAGTGTG 3106

RESULT 13  
 CVMEM11ZP  
 LOCUS CVMEM11ZP 3221 bp DNA circular SYN 05-JUL-1999  
 DEFINITION Cloning vector pGEM-11Zf(+).  
 ACCESSION X65313  
 VERSION X65313.2 GI:5420374  
 KEYWORDS beta-lactamase; bla gene; cloning vector; lacZ gene; multiple cloning site; phase I1 region; promoter; pUC/M13 primer binding site.

SOURCE  
 ORGANISM Cloning vector pGEM-11Zf(+).  
 artificial sequence; vectors.

REFERENCE  
 AUTHORS 1 (bases 1 to 3221)  
 TITLE Technical Services.  
 JOURNAL Direct Submission  
 Submitted (23-MAR-1992) Technical Services, Promega Corporation,  
 2800 Woods Hollow Road, Madison, WI 53711-5399, USA

REMARK  
 REFERENCE 2 (bases 1 to 3221)  
 AUTHORS Technical Services.  
 TITLE Direct Submission  
 JOURNAL Submitted (05-JUL-1999) Technical Services, Promega Corporation,  
 2800 Woods Hollow Road, Madison, WI 53711-5399, USA

COMMENT  
 On Jul 8, 1999 this sequence version replaced g1:58168.  
 See X65300-X65335 for related vector sequences  
 This vector can be obtained from Promega Corporation, Madison, WI.  
 Call one of the following numbers for order or technical information:  
 Order or Technical 800-356-9526  
 In Wisconsin 800-356-9526  
 Outside U.S. 608-274-4330.

## FEATURES

## source

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature



```

LR5ALPAGWFIADKSGSGEGRSGRIIALALDPDGKPSRIIVITYITGSGATMDERRKQIN
EIGASLIRKH"
2586. .3041
misc_feature
/note="phage fl region"
3042. .3202
misc_feature
/note="lac operon"
3162. .3178
misc_feature
/note="pUC/M13 forward primer binding site"
3205. .3221
promoter
/note="T7 promoter"
BASE COUNT      826 a      793 c      825 g      777 t
ORIGIN

```

|                            |       |               |          |             |
|----------------------------|-------|---------------|----------|-------------|
| Query Match                | 35.4% | Score 2809.4  | DB 56    | Length 3221 |
| Best Local Similarity      | 99.5% | Pred. No. 0   |          |             |
| Matches 2630; Conservative | 0     | Mismatches 11 | Indels 4 | Gaps 1      |

|    |      |  |      |
|----|------|--|------|
| QY | 5091 | ccctctcattatagatcgccaaacgcgcggaggaagcggttgctgctatctggcgctc         | 5150 |
| Db | 290  | CCACCTCATTAATTAAGATCGCCCAACGCCCGGGAGAGCGGTGGCTATTGGGCGCTC          | 349  |
| QY | 5151 | ttccgcctccctcgcctcaactbactcgcctgcctgcgtgcgttcgtctgcctgcgcgcagcgatc | 5210 |
| Db | 350  | TTCCGCTTCCCTCGCTCACTGACCTCGCTGGCGTGGCTGTTCGCTCGCGGACGCGATTC        | 409  |
| QY | 5211 | agctcactcaaaagcggtlaatacgyttatccacaagaatcaagggtlaacgcaggaanaa      | 5270 |
| Db | 410  | AGCTCACTCAAAAGCGGTATACGGTTATCCACACATACAGGGGTAAACGACGAAAGAA         | 469  |
| QY | 5271 | catgtgacaaaagcgcaacaaaggccagagacccgttaaaggccggtttgcctgcgcgtc       | 5330 |
| Db | 470  | CATTGTGCAAAAGCCAGACAAAAGGCCAGGAACCGTAAAGGCCGCGTGTGCGGCTT           | 529  |
| QY | 5331 | tttcataagctccgcgcgcgccttcagcagatatacaaaaatcagcctcaagtcaagctg       | 5390 |
| Db | 530  | TTTTCATATAGGCTCGGCCCGCCCTTGACGAGCATCAAAAATGACGCTCAAGTCAGAGGTG      | 589  |
| QY | 5391 | gcgaaacccgacacagagactataaagataccagcgctttcccccctgtaagctccctcgtgcg   | 5450 |
| Db | 590  | GCGAACCCGACAGGACTATAAAGATACCAAGGCGTTTCCCTCGAAGCTCCCTCGAGCG         | 649  |
| QY | 5451 | ctctccctgtctcgaacctgcgccttaccggaatacctctgcgcgccttcccttccttgggaag   | 5510 |
| Db | 650  | CTCTCCGTGTTCGACCCCTGCGCTTACGCGGATACCTGTCGCCCTTCTCCCTTGGGGAG        | 709  |
| QY | 5511 | cgtgcgccttctctcaatbgtcacaacgcgtgtagatctcagttcgtgtgtagctgtctc       | 5570 |
| Db | 710  | CGTGCCGCTTTCATAGCTACAGCGTGTAGTATCATCAAGTTGGGTGAGTGTTCGCTC          | 769  |
| QY | 5571 | caagcttgagctgtgtgcacgaaaccccccggtttagcccgacacgcctgcgccttaccggtaa   | 5630 |
| Db | 770  | CAAGCTGGGTGTGTGCAGAACCCGCCCTTACAGGCCACGCTCGCTTATCCGGTAA            | 829  |
| QY | 5631 | ctatgcgtcttgaatctcaacccggtlaagaacagactatcgccactgcgcgcagcagcacgtg   | 5690 |
| Db | 830  | CTATCGCTTGAATCCAACCCCGGTAAAGACACACTTATTCGCCACTGCGACGACCACTCG       | 889  |
| QY | 5691 | taacagatagcagagcgaagtatgtlaagcgtgtcacaagatcttctlaagtgtgtgcc        | 5750 |
| Db | 890  | TAAACAGATTAGCAGAGCGAGGTATGTAGCGGCTGTACAGAGTTCTTGAAGTGTGTGCC        | 949  |
| QY | 5751 | taactcgcgttaactaagaagaacgattttgttatctgcgcctctgcctggaagccaagttaac   | 5810 |
| Db | 950  | TAACTACGGCTTACACTAGAAAGAACATATTGTGTCTGCGCTCGCTGAAGCCAGTTTAC        | 1009 |
| QY | 5811 | cttcggaanaaagctgtgttagctctctgtatccggcaacaacaacacgcgtgtgttagcgttgc  | 5870 |
| Db | 1010 | CTTCGGAAAAAGAGTGTGTGATGCTCTTGATCGCGCAAAACAACACCGCTGTACGGGTGC       | 1069 |
| QY | 5871 | tttttttcttgctcaagcagcagattacgcgcgagaanaaaagatctcaagaagctctt        | 5930 |
| Db | 1070 | TTTTTTTTGTTTGCACACACACAGTTTACGGCGAAGAAAAAGAGTCTCAAGAATCTCTTT       | 1129 |

|    |      |   |      |
|----|------|---|------|
| QY | 5931 | gaccttctctacggggtctgaagctcagdygaacgcgaaactcaagttlaagggatttgyt       | 5930 |
| Db | 1130 | gATCTTTCCTAGGGGGGTGTGACGCTCAGTGGAAAGAAACTACGTTAAGGGATTTCGT          | 1189 |
| QY | 5991 | catgagattaccaaaggatcttcacctagatccctttaaattaaatgaagtttaa             | 6050 |
| Db | 1190 | CATGAGATTATCAAAAAGATCTTCACCTTGATCCCTTTAAATTAAATGAAGTTTAA            | 1249 |
| QY | 6051 | atcaacttaagtatataatgaagtaaaccttgcctgacgttaccatgcttcaatcagtga        | 6110 |
| Db | 1250 | ATCATCTTAAAGTGTAAATATGATGAATACCTTGCTGACAGTTACCAATGCTTAATCAGTA       | 1309 |
| QY | 6111 | ggacacctctccacgggtctgctcatcttctgtatccatgaagtgtcgtagctcccgctgt       | 6170 |
| Db | 1310 | GGCACCTATCTCACAGATCTGTCTATTTCGTATCCATAGTGGCTGTGACTCCCGCTGT          | 1369 |
| QY | 6171 | gtagaataactacgaatacggggaaggcttaccatcttgccccagtgctgcgaatgataccg      | 6230 |
| Db | 1370 | GTAGATTACCTACGATACGATGCGGGGCTTACATCTGGCCCCAGTGGCAATGATACGCG         | 1429 |
| QY | 6231 | agaccacagctcaccggtctccagattatcagcaataaacccagccagccgaaggccga         | 6290 |
| Db | 1430 | AGACCCACAGCTCACCGGCTCCAGATTTATCAGCAATAAACACAGCCAGCGGAAGGGCCGA       | 1489 |
| QY | 6291 | gcgcgaagatgtctctgcacattatccgcctccatccagatccatataatgtctgcggga        | 6350 |
| Db | 1490 | GCGCAGAAAGTGGTCTCTCAACTTTATCCGCTTCATCCAGTCTATTAAATTGTTGCCGGGA       | 1549 |
| QY | 6351 | agctagagtaagtagtctgcgcagttaaagtctgcgcaagcttctgcccattgctacag         | 6410 |
| Db | 1550 | AGCTAGAGTAAAGTGTTCGCGCAGATTAATGTTGCCCAAGTGTGTGCATGTCTACAG           | 1609 |
| QY | 6411 | catcgttgtctcaagctcctgccttttgtagtgcctatccagctccogtctcccaagctc        | 6470 |
| Db | 1610 | CATGCTGGTGTCAAGGCTCGTGGTTGGTATAGGCTTATTAAGCTCCGGTTCCTCAACGATC       | 1668 |
| QY | 6471 | aaggcgagttacacgatcccccacttctgtgcgcaaaaagcggtttagctccctgcgtccc       | 6530 |
| Db | 1670 | AAGCGAGTTACATGATCCCCCATGTTGTCAMAAAAAGGGTTAGCTCTTGGTCTCC             | 1729 |
| QY | 6531 | gatcgttgtcagaagaatgaagtgtgcgcgagtggtataccatcagtggtatbgcagcactga     | 6590 |
| Db | 1730 | GATCGTGTGTCAAGTAAAGTTGGCGGAGGTTATACATGATGTTATGGCAGCACTCCA           | 1789 |
| QY | 6591 | taattccttactctcatgcccacgcgtaagaabtccttctcgtacgtggtgaatactaac        | 6650 |
| Db | 1790 | TAAATTCTTACTACTCATGCGCATCCGTAAGATCCTTTCTGTGACTGGTGAATACTAAC         | 1849 |
| QY | 6651 | caagtcattctgagaatagtgtaatgcgcgcgaccgagttgcctctgcgccgcgtcaatag       | 6710 |
| Db | 1850 | CAAGTCATTTCTGAAATAGTATGCGGCGACCGAATTCCTCTTGCCGCGCTCAATACG           | 1909 |
| QY | 6711 | ggataataccggcgccacataagcagaactttaaagtgctcatcatcttgtaaaaagcttttc     | 6770 |
| Db | 1910 | GGATTAATACCGCGCCACATATGCGAAGACTTTAAAGTGCATCATTTGGAAACGTTCTTC        | 1969 |
| QY | 6771 | ggggcgaaaaactctccaaagatactctacgcgtgtgtgaatcagttcagatgtaaacccacg     | 6830 |
| Db | 1970 | GGGGCGAAAACOTCTCAAGGATCTTACCGGCTTGAGATCCAGTTCGATGTAAACCATCG         | 2029 |
| QY | 6831 | tgcacccaacttgatctctcagcatccttcttactctacccagcgcttctcgtgtgacaaaac     | 6890 |
| Db | 2030 | TGCACCCAACTGATCTTCACACTCTTTTACTTTTCACACACGCTTTCGGGTGCGCAAAAAC       | 2089 |
| QY | 6891 | aggaagaacaaatgycgcgcaaaaaaggaataagcggcagacaggaatcttgataactcat       | 6950 |
| Db | 2090 | AGGAAGCGAAATATGCCCCAAAAAAGGATTAAGCGGACAGCGAATGTGAATCTCAT            | 2149 |
| QY | 6951 | acctctccctttccaatatattgaagcatcttacaaggttatcaagggtatgctcotaatgacgata | 7010 |
| Db | 2150 | ACTCTTCCTTTTCAATATATATGAAGCAATTAATACAGGGTATATGCTCATACAGCGGATTA      | 2209 |

QY 7011 catatttgaaatglatctagaacaaatacaaaatagagggttcgcgcacatactcccccga 7070  
|||||  
Db 2210 CATATTGAAATGATATTGTAAGAAAAATAAACAAATAGGGGTTCGCCACATTTCCCGGAAA 2269  
QY 7071 agtgcaccactgagctctagaacacatatactatgacatcaactaataaataagagg 7130  
|||||  
Db 2270 AGTGCACCTGACGCTATGAGAAACCATTTATATCATGACATTAACCTATTAATAATAGGG 2329  
QY 7131 tacaagagggcccttcgtctcgcgcgtttcgtgatgacaggtgaanaaccttgacacat 7190  
|||||  
Db 2330 TATACAGAGGGCCCTTTCGCTCGCGCGTTTCGTCATGACGGTGAAACCTCTGACACAT 2389  
QY 7191 gcagctcccggaagacggttcaacagcttctgtlaagcggatgcgggagcagaacggccg 7250  
|||||  
Db 2390 GCACCTCCCGGAGAGCGCTACAGAGCTTGTCTGTAAACCGGATGCGGGAGCAGCAAGCCCG 2449  
QY 7251 tcaagggcgctgcagcgggttgttgcggggtgtcggggcttgactaactaagcgatcaga 7310  
|||||  
Db 2450 TCAGGGCCCGTACGGGGGTGGGGGTGGGGGTGGGGGTGGGGTGAACCTATGCGGCATCAGA 2509  
QY 7311 gcagattgactgagagtgcaacacatgctggtgtgaataacccgcaacagatgcgtlaagag 7370  
|||||  
Db 2510 GCAATTTCTACTGAGAGTGCACACATATCGGTGTAATACCGCACAGATGCGTAAAGAG 2569  
QY 7371 aaaaaccgcatcaggaaattgtlaaagctlaatatctgtlaaataatcgcgtlaaattt 7430  
|||||  
Db 2570 AAAATACCGCATCAGGAATTTGTAAAGCTTAATATTGTTAAATTCGCGTTAAATTTT 2629  
QY 7431 tgtaaatcacctcatatttttaacaaatagggcaaatcggaaataatccctataatca 7490  
|||||  
Db 2630 TGTTAATCACCTCATTTTATTAAACCAATAGCGCGAATCGGCAAAATCCCTTATTAATCA 2689  
QY 7491 aaagaatagacggagataggttgtgagtgctgtccagttctggaacaaagatccactatla 7550  
|||||  
Db 2690 AAAGAAATGACCGAGATAGGTTGAGTGTGTTCCAGTTTGAACAAAGATGCACATATTA 2749  
QY 7551 aagaagctgagctcccaagctcaaaaggcgaaaaaccgtctcatcaggcgagatgcccacta 7610  
|||||  
Db 2750 AAGAACGTGAGCTCCAAAGCTCAAAAGGGGCAAAACCGTCTATCAGGGCGCATGGCCACTA 2809  
QY 7611 cgttaaccacacaccccttaactaagtttcttgaggctgaggtgctgcgtlaaagcactaactcg 7670  
|||||  
Db 2810 CGTAAACCATATACCTTAATCAAGTTTTCGGGGTTCGAGGTGCCGTAAACACTTAATTCGG 2869  
QY 7671 aaccctaaaggagcccccgatcttagaagcttgaacggggaaaagccggcgagacgttgacgaga 7730  
|||||  
Db 2870 AACCTTAAGGAGAGGCCCGCGATTGAGAGCTTGACGGGGAACCGCGCAACCTGGGGAGA 2929  
QY 7731 aaggaaagggaagaaagcgaaaggagcggtgcttagaggcgcttggaagtgtagcggtcacg 7790  
|||||  
Db 2930 AAGGAAGGGAAGAAAGCAAAAGGAGCGGGCGCTTAGGGCGCTTGCAAGTGTACCGGTCAAG 2989  
QY 7791 ctggcggttaaccacacaccccgcgctcttaatgagcgctcaagggcggtcgcgccat 7850  
|||||  
Db 2990 CTGGCGGTAAACCATCACACCGCGCGCTTAATGCGCGCTTAACAGAGCGCGT-----CCAT 3045  
QY 7851 tcgcatctcagggtctacgaaactgtctggaagggcgatcggtgcgggccccctctcgtcatla 7910  
|||||  
Db 3046 TCGCATTCAGAGCTGCGCAACGTTGGGAAGGGGCGATCGGTGCGGCGCTTCCTCGCATTA 3105  
QY 7911 cgccaagctggtctgcagggggggggg 7935  
|||||  
Db 3106 CGCCAGCTGCGCAAAAGGGGATGTG 3130

ORGANISM unidentified  
REFERENCE unclassified.  
AUTHORS 1 (bases 1 to 4118)  
TITLE Dahm, M.W.  
JOURNAL METHOD OF QUANTIFYING TUMOUR CELLS IN A BODY FLUID AND A SUITABLE  
TEST KIT  
PATENT: WO 9718322-A 17 22-MAY-1997;  
DAHM MICHAEL W (DE)  
FEATURES Location/Qualifiers  
source 1..4118  
/organism="unidentified"  
/db\_xref="taxon:32644"  
BASE COUNT 991 a 1032 c 1094 g 1001 t  
ORIGIN  
Query Match 35.2%; Score 2794.2; DB 9; Length 4118;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 2828; Conservative 0; Mismatches 13; Indels 6; Gaps 2;  
QY 5091 ccagctgcatlaatgaatcgcgcgaacgcgcggggaagcggttgcgtatctggcgctc 5150  
|||||  
Db 1185 CCAGCTGCATTATATGATACGCGCAACGCGGGGAGAGCGGTTTGGTATTTGGCGCTC 1244  
QY 5151 ttccgcttcctcgtcactgactgcgtgcgtcgggttcggtcgtgcggcgagctatc 5210  
|||||  
Db 1245 TTCCGCTTCCTCGTGCACGTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTATC 1304  
QY 5211 agctcaactcaaaaggcggaatagcgttlaacaaataacaggggaataacggaagaa 5270  
|||||  
Db 1305 ACCTCACTCAAAAGCGGGAATACGGTTATCCACAAATCAAGGGATACGCAAGAAAGAA 1364  
QY 5271 catgtgagcaaaaggccagcaaaaggccaagaaacggtlaaaaggccggttgcgtgcgt 5330  
|||||  
Db 1365 CATGTGAGCAAAAGGCCAGCAAAAGCCAGGAACCGTAAAGGGCGGCTTGCGGCTT 1424  
QY 5331 ttccatagagctcgc 5390  
|||||  
Db 1425 TTTCATATAGGCTCCGCCCGCTGAGGACATCACAAATGAGCGCTCAAGTACAGAGTG 1484  
QY 5391 gggaaaccggaagacactaagaatacaccagcggttcccccctggaagctccctgcgtcg 5450  
|||||  
Db 1485 GCGAAACCCGACAGAGACTATTAAGTATACCAAGCGTTTCCCTTGGAAAGCTCCCTCGTGC 1544  
QY 5451 ctctcgttcccgaccctgcgcgtctaccggaatacctgcgcgttctcccttcgggaag 5510  
|||||  
Db 1545 CTCCTCGTTCGACCGCGCGCTTACCGGATACCGTACCGCTTTCCTTCCCTTGGGGAAG 1604  
QY 5511 cgtgcgcgttctcaagtgctcaagctgaggtatcctcaagttcgtgtgaggtcgtctc 5570  
|||||  
Db 1605 CCGTGCCTTCTCATATACCTCACGCTGTAGGTATCTCAGTTCGGGTGAAGTCTGCTTC 1664  
QY 5571 caagctggcggtgtgtgcacgaaccccccgcttcagcccgcaacgcgtgcctatccgtaa 5630  
|||||  
Db 1665 CAAGCTGGGCTGTGTGACGAACCCCCGTTTCAGCCGACCGCTGCGCTTATCCGGTAA 1724  
QY 5631 ctatcgtcttgagtcacaaaccggtlaagaacagacttaccgacactgcgaagcagccactg 5690  
|||||  
Db 1725 CTATCGTCTGAGTGCACACCGGTAAAGACAGACTTATCCGCACTGGAGCAGCACACTGG 1784  
QY 5691 taacagagattagcagagcaggtatgtagcggtgtcagagatctctgaagtggtggc 5750  
|||||  
Db 1785 TAACAGGATTTAGCAGAGGAGGTATGTGGGCTGTCAAGATTTCTTGAAGTGTGTGGCC 1844  
QY 5751 taactagcgtacacttagaaggaagtaatttgatctcgtcgtcgtcgtcgtcgtcgtcgt 5810  
|||||  
Db 1845 TAACTACGGCTACACTTAAGAGCAGTATTTGTATCTGCGCTTGTGTAAGCAAGTTTAC 1904  
QY 5811 ctctggaagaaaggtctggttagctcttgaatccggcaacaaacacacccgtgtgtagcgttg 5870  
|||||  
Db 1905 CTTGGGAAAGAGTTGGTAGTCTTGAATCGGCAAAACCAACCGCTGTGTACGGGTGG 1964  
QY 5871 ttltttgttgtaagcgacagattagcgcgcaaaaaaaggatctcaagaagatccctt 5930

|||||  
Db 1965 TTTTGTGTCAGAGCAGATTACGCCGAGAAAAAGATCTCAAGAGATCTTT 2024  
Qy 5931 gatctttctcagggtcgtcagtcagtggaagaaaacacagtaagatttggt 5930  
Db 2025 GATCTTTCAGGGGTCGTGAGCTCACTGAGTGAAGAAAACACAGTTAAAGGATTTTGGT 2084  
Qy 5991 catgagatatacaaaaaggatctcacccatagatcctttaaataaalaagaatttaa 6050  
Db 2085 CATGAGATTATCAAAAAGAGATCTTACCTAGATCTTTAAATTAAAAATGAAGTTTAA 2144  
Qy 6051 atcaatctaaagiatatataagaaacttggctcagagttacaaatgcttaacagtga 6110  
Db 2145 ATCATCTTAAAGTATATATAGTAACTGGCTGACAGATTACCAATAGCTTAATCAGTGA 2204  
Qy 6111 ggcacatctcccaagcagatctgtctatttcgtctcaatagttgcccagaccctcgtcgt 6170  
Db 2205 GGCACCTATCTCACCGAGTCTCTATTTTCATTTCATCCATAGTTGCTGACCTCCCGTCTGT 2254  
Qy 6171 gtagataactacgatacagagagagcttaaccactcggcccaagtcgcaatgataccgag 6230  
Db 2265 GTAAATTAACCTAGATAGGAGGGGCTTACCACTTGGCCCGCTGCTGCAATGATACCGCG 2324  
Qy 6231 agaccacagctcaccaggtcccaagattatcaagcaataaaccagccagccggaagagcgga 6290  
Db 2325 AGACCCACAGCTCACCGGCTCCAGATTATACGAATTAACACGACGACCGGAGGCGCGA 2384  
Qy 6291 ggcgaagaagtggtctcgaacttatccgcgtccatccagttctataatggttcgagga 6350  
Db 2385 GCGGAGAGTGTGCTCTCACTTATCCGCTCCATCCAGCTATTAATTTGTTCCCGGGA 2444  
Qy 6351 agctagaagtaagtagtcgcagcttaagtttgcgcaagctgttcgcatgtgtacag 6410  
Db 2445 AGCTAGAGTAGTAGTGTGCGCAGTTATAGTTTGGCGCAAGCTTGTCCATGTGTAAACG 2504  
Qy 6411 catcgtgtgtcagcgtcgtcgttggatagcttcaatcgaagtcgggttcccaagcagtc 6470  
Db 2505 CATGCTGTGTCAACGCTCGTGTGTGATGAGCTTCAATTCAGCTCCGGTATCCCAACGATC 2564  
Qy 6471 aagcgaggttaacatgatacccccattgtgttcaaaaagcggtgagcttgcgtcc 6530  
Db 2565 AAGCGAGTTTACATGATGCCCATGTTGTGCAAAAAGCGGTAGCTCTTCCGTTCTCTCC 2624  
Qy 6531 gatcgtgtcagaagtaagtttgcgcagtggttatacctaagtttgaagcagctgca 6590  
Db 2625 GATGTTGTGTGAGTAGTATGTCGCGCAGTGTATTCATCTAGTGTATGGAGACACTGCA 2664  
Qy 6591 taattctctactgtcaatgcacatcgtlaagatgctttctcgtgagcgtgtgaatcgaac 6650  
Db 2685 TAAATCTCTTACTGTCAATGCCATCCGTAAGATGCTTTCTGTGACTGCTGATCTCAAC 2744  
Qy 6651 caaatcatctcgaagaatgtagtgcgagcgaagtgctgtcccgcgcgtaacag 6710  
Db 2745 CAATCTATCTTGAAATATGTTATGCGCGACCGAGTTGCTTGGCCGGGCTTAATAACG 2804  
Qy 6711 ggaataatacgcgcacataagcagaacttlaaagtgctcatcaatitgnaaacgtttctc 6770  
Db 2805 GGATTAATACCGCGCCACATAGCAAGACTTTAAAGTGTCTCATCTTGGAAAAAGCTTTCTC 2864  
Qy 6771 gggcggaagaactctcaagatcttaacgctgttgagaatcgaatgtaaccactcg 6830  
Db 2865 GGGCGCAAAACTCTCAAGAGATCTTACCGCTGTGAGATCCAGTTGTAACCACTGTCG 2924  
Qy 6831 tgcacccaactgactctcaacatcttacttcaacgaaggttctcgtggaagaaaaac 6890  
Db 2925 TGCACCCAACTGATCTTCAGCATTTTACTTTTACCAAGGCTTCTGGGTGAGCAAAAAC 2984  
Qy 6891 aggaagcgaaatagccgcaaaaaggaataaaggcgacacggaatgctgaatactcat 6950  
Db 2985 AGGAAGCGAAATAGCCGCAAAAAGGAATTAAGGCGCACACGGAATGTTGAATACATCAT 3044  
Qy 6951 actctcccttccaataataatgaagcaattalcaaggttatgtctcagcagcgata 7010  
|||||

Db 3045 ACTCTCTCTTTTCATATATTATGAGCAATTTATCAGGGTATTGTCTCATGAGCGGATA 3104  
Qy 7011 catattgaagtagtatagaaaaaaataacaatatgggttccgcgcattccccgaaa 7070  
Db 3105 CATATTTGAATGTATTTAGAAAAATAACAATATGGGTTCCCGCACATTTTCCCGCAAA 3164  
Qy 7071 agtccacactgagcttcaagaacacatattatatcatatcacatlaaacctataaaatagcg 7130  
Db 3165 AGTCCACACTGACGCTCTTAAGAAACATTATTATCATGACATTAACCTATAAAATTAAGCG 3224  
Qy 7131 tatacgaaggcccttcgtctcgcggttccgtgtatgacgtgtgaaaaacbtctgacacat 7190  
Db 3225 TATCAGCAGGCGCTTTCGTCTCGCGCTTTCGTGATGACGAGTGAACCTCTGACACAT 3284  
Qy 7191 gcaagctcccggaagcagttcacagctgtctgtgaagcggaatgcggggagcagacaagccg 7250  
Db 3285 GCACGCTCCGGAGACGGTCAACAGCTTGTGTAAAGCGGATGCGGGAGCAGACAAAGCCG 3344  
Qy 7251 tcaaggcgctcgaagcggtgttgagcggtgtcggggctgtgaactatgcgacatcaga 7310  
Db 3345 TCAGAGGCGCTCACGGGTTGTGGCGGTTGTGGGGCTGGCTTAACATATGGGCATCAGA 3404  
Qy 7311 gcaagattgtaactgagagtgcaacaaatagcggtgtgaaataacgcagacagatgctgaaggag 7370  
Db 3405 GCAGATTGTACTGTGAGAGTGCAACCATATGCGGTGAATACCCGACAGATGCTTAAGGAG 3464  
Qy 7371 aaaaataccgatacag--gaatgttaagttataatttgtaaaatttcggttaatt 7428  
Db 3465 AAAATACCGCATACAGCGGAATGTGAACGTTATATTTGTTAAATTCGCGTTAAATA 3524  
Qy 7429 ttgttaaatcaagctcaatttllaaccaataagcggaataatcgcaaatccctataat 7488  
Db 3525 TTTGTTAAATCAGCTCATTTTTTTAAACCAATAGCGCAATCGCAAAATCCCTTATTAAT 3584  
Qy 7489 caaagaatagaacggagataaggtgtgagtggttccagtttggacaagaagttccactat 7548  
Db 3585 CAAAAGATATGACCGAGATAGGGTGTGAGTGTGTCCAGTTTGGAACAAGATGCTCACTAT 3644  
Qy 7549 taaagaacgttgagctccaagctcaaaaggcgaaaacggtctacagggcgatgtgcccac 7608  
Db 3645 TAAAGAACGTGAGCTCCACGCTCAAGAGGCGAANAACGCTATACAGGCGATGCCCCAC 3704  
Qy 7609 taactgaacacatacccttaactcaagtttcttggtgtcaggtgtgcgttaaacataatc 7668  
Db 3705 TAGCTGAACCATACCCCAATTAAGTTTTCGCGTGCAGTGCCTTAAGCTCTAATATC 3764  
Qy 7669 ggaacccctaaaggagcccccagatttagagcttgaacggggaagccggcgaaogtgcga 7728  
Db 3765 GGAACCTTAAGGGAGGCCCGCGATTTAGAGCTTGACGGGGAAGCCCGCAACGCTGGCGA 3824  
Qy 7729 gaaaggaagggaagaagcgaaaggagcgagcgctlaagggcgtggaagtgtaggggtca 7788  
Db 3825 GAAAGGAAGGGAAGAAAGCAAGAAAGAGAGGGGCGCTAAGGGCGTGAAGTGTACCGGTCA 3884  
Qy 7789 cgtctgcgttcaaccacacacacccgcgcgttlaatgagcgcgcttacaaggtgcgtcgccg 7848  
Db 3885 CGCTGGCGGTAAACACACACACCGCGCGCTTAATGCGCGGTAAACAGAGGGCGGT---CC 3940  
Qy 7849 attcgccattcaaggtctacgcgaactgttgggaaggcggaatcgtgtgcggtctctgctat 7908  
Db 3941 ATTTCGCAATTCAGGCTGCGCAACGTGTGGGAAGGCGATGCTGCGGGCTCTTCCTAT 4000  
Qy 7909 taagcagcgtgcgtgcagggggggggg 7935  
Db 4001 TACGCCAGCTGGCGAAAGGGGATGTG 4027  
|||||

RESULT 15  
A93964 LOCUS A93964 4118 bp DNA  
DEFINITION Sequence 18 from Patent WO9718322.  
ACCESSION A93964  
VERSION A93964.1 GI:6742065

[illegible]

|    |      |   |      |
|----|------|---|------|
| QY | 5871 | tttttttcttgccaagcagcagatctacgcgcagaaaaaagatctcaagaagatcctt        | 5930 |
| Db | 1965 | ttttttttgtttggcaagcagcagatttacgccacaaaaaaagatctcaagaagatcttt      | 2024 |
| QY | 5931 | gattcttctcagggggtctgaagcttagtbgaaacgaataactcgtttaaggaatttbg       | 5990 |
| Db | 2025 | gattcttttctcagggggtctgaagcttagtbgaaacgaataactcagcttaagggaattttt   | 6084 |
| QY | 5991 | catgagatctatcaaaaagatctctacctaagatcctttaataaaatgaagitttaa         | 6050 |
| Db | 2085 | catgagattatcaaaaaagatctttcaccttgatctcttttaaaatgaagctttttaa        | 2144 |
| QY | 6051 | atcaatctaaagtctataitagaatnaacttgccttgcaatttcacaaatggttaacag       | 6110 |
| Db | 2145 | atcaatctaaagtctataitagaatnaacttgccttgcaatggttaacaaatggttaacag     | 2204 |
| QY | 6111 | ggcacctctcccaagcagatctgtcattctgtctcatcaagtgtgcctgaccccgctgt       | 6170 |
| Db | 2205 | ggcacctctctcagcagatctgtcattctgtctcatcaagtgtgcctgaccccgctgt        | 2264 |
| QY | 6171 | gtagaataactcagatcacgggaagggtctacaatcttgccccagtgctgcaatgaacgcg     | 6230 |
| Db | 2265 | gtagaataactcagatcacgggaagggtctacaatcttgccccagtgctgcaatgaacgcg     | 2324 |
| QY | 6231 | agacccacgcctcacacgggtctcagaattatcaagaataaaacacagccagccggaaggccga  | 6290 |
| Db | 2325 | agacccacgcctcacacgggtctcagaattatcaagaataaaacacagccagccggaaggccga  | 2384 |
| QY | 6291 | gcgcagaagtggtctctgaacttatcccttcacatccacagctbatnaattgttgcggga      | 6350 |
| Db | 2385 | gcgcagaagtggtctctgaacttatcccttcacatccacagctbatnaattgttgcggga      | 2444 |
| QY | 6351 | agctagaataagtagttcgcgaagttaaagtgttgccgaacgtttgtgcattgtcaagg       | 6410 |
| Db | 2445 | agctagaataagtagttcgcgaagttaaagtgttgccgaacgtttgtgcattgtcaagg       | 2504 |
| QY | 6411 | catcgttggtgcacgcctgcgtctgttggtaagcttcattcaagctccggtctcccaagatc    | 6470 |
| Db | 2505 | catcgttggtgcacgcctgcgtctgttggtaagcttcattcaagctccggtctcccaagatc    | 2564 |
| QY | 6471 | aagcgcagatctacagatcccccagctgtgtgtcaaaaaagcgttagtctcctctggtcccc    | 6530 |
| Db | 2565 | aagcgcagatctacagatcccccagctgtgtgtgtcaaaaaagcgttagtctcctctggtcccc  | 2624 |
| QY | 6531 | gacgcttgctcgaagtaagtgtgcccagtgcttatcaactcagatggtatagcagacgtca     | 6590 |
| Db | 2625 | gacgcttgctcgaagtaagtgtgcccagtgcttatcaactcagatggtatagcagacgtca     | 2684 |
| QY | 6591 | taattctcttaactgctcatgcatccatccgtlaagaatgctttctgtgaactggtgaatctaac | 6650 |
| Db | 2685 | taattctcttaactgctcatgcatccatccgtlaagaatgctttctgtgaactggtgaatctaac | 2744 |
| QY | 6651 | caaatcaatctcgaagaataagtgtatgcggcgaacagagttgcctcttgccggcgctcaaatag | 6710 |
| Db | 2745 | caaatcaatctcgaagaataagtgtatgcggcgaacagagttgcctcttgccggcgctcaaatag | 2804 |
| QY | 6711 | ggaataataccgcgcacataagcagaactttaaagtgctcatcatcttgaaaaagcttcttc    | 6770 |
| Db | 2805 | ggaataataccgcgcacataagcagaactttaaagtgctcatcatcttgaaaaagcttcttc    | 2884 |
| QY | 6771 | ggggcggaataactctcaaggaatcttaacgcgtgttgagatccagttcgaatgaaccacatcg  | 6830 |
| Db | 2865 | ggggcggaataactctcaaggaatcttaacgcgtgttgagatccagttcgaatgaaccacatcg  | 2924 |
| QY | 6831 | tgcaaccaactcagatcttcaagatcttcttaacttccacaagcgtttctcgggtgagcaaaac  | 6890 |
| Db | 2925 | tgcaaccaactcagatcttcaagatcttcttaacttccacaagcgtttctcgggtgagcaaaac  | 2984 |
| QY | 6891 | aggaagaagcaaatgycgcgaaaaaagggaataaggcgcagacaggaatglttgaatctcat    | 6950 |
| Db | 2985 | aggaagaagcaaatgcccccaaaaaagggaataaggcgcagacaggaatggtgaatctcat     | 3044 |

QY 6951 acctcccttttcaatattgaagcattltaaggttatgtctcaatgaagcgata 7010  
|||||  
Db 3045 ACCTCTTCCTTTTCAATATTATTAACCATTTTATCAGGGTTATTGTCTCATGAGCGSATA 3104  
|||||  
QY 7011 cataattgaaatgtaattagaanaaaataacaataaggggttcgcgcacattccccgaaa 7070  
|||||  
Db 3105 CATATTGAAATGATTTAGAAAAATAAACAATAGGGGTTCCGGGCACATTTCCCGGAAA 3164  
|||||  
QY 7071 agtgcacccctgaagctcctaagaacattatataatgaacttaacctataaataagcg 7130  
|||||  
Db 3165 AGTCCACCTGACGCTTAAGAAACCATTTATCATATGACATTAACCTTAATAAATAGGCG 3224  
|||||  
QY 7131 tatacgaaggcccttcgtctgcgcgttcgcgtgaacggtgaacacctgaacat 7190  
|||||  
Db 3225 TATACAGAGGCCCTTGTGTCTCGGGCTTTCGGTGATGACGGTGAANAACCTTGACACAT 3284  
|||||  
QY 7191 gcagctcccgagagacggtlcaacagctgtctgaagcgaatgcgcggagacgaacagccg 7250  
|||||  
Db 3285 GCAGCTCCCGAGACGGGTCAACAGCTGTGTAAAGCGGATGCCGGAGACAGACGCCG 3344  
|||||  
QY 7251 tcaggcgcgctcagcggtgtgtggcggtgtgcgggttgtaactatgcgcacatga 7310  
|||||  
Db 3345 TCAGGGCGGCTGACGGGCTGTGGGGGTGGGGCTGCTTAACATATGCGGCATCAGA 3404  
|||||  
QY 7311 gcagattgtaactgaagctgcacataatgcgtgtgaataacgcacacagatgcgtaagag 7370  
|||||  
Db 3405 GCAGATTGTCTGAGAGTGCACCATATGCGGTGTGAATATACCGCACAGATGCGTAAGGAG 3464  
|||||  
QY 7371 aaaaaccgcatcag--gaaatgtaaacgtlaatatlttgtlaaaatlccgcgtlaaat 7428  
|||||  
Db 3465 AAATACCGCATCAGGCGAATGTAAACGTATATTTTGTAAATTCGCGTTAATA 3524  
|||||  
QY 7429 ttgtttaaatacagctaatltttaaaccaatagccggaatccgcgaataatccctataat 7488  
|||||  
Db 3525 TTTGTAAATCAGCTCATTTTAAACCAATAGGCGCAATCGCAAAATCCCTTATTAAT 3584  
|||||  
QY 7489 caaagaatagaccgaatagaggtltgagltgtlccaglttgaacaagatccactat 7548  
|||||  
Db 3585 CAAAAGATATGACCGAGATAGGGTGTAGTGTGTTCAGTTTGGACAAAGTCCACTAT 3644  
|||||  
QY 7549 taaagaacgttgaactccaacgtcaaaagcgcaaaaacgltcatalcaaggcgatggccac 7608  
|||||  
Db 3645 TAAAGAACGTGACTCCAAAGTCAAAAGGCGAAAAACCGTCTATCAGGCGGATGCCAC 3704  
|||||  
QY 7609 taagtgaaccatcacccatacaagtttttgggtgcaggtgcgtaaaagcaactaatc 7668  
|||||  
Db 3705 TACGTGAACCATCAACCAATCAAGTTTTCGCGTCAAGGTGCGCTTAAGCTCTTAATC 3764  
|||||  
QY 7669 ggaaccctlaaaggagcccccgaatttagagcttgaacggtgaagccgcgaacgtgagcga 7728  
|||||  
Db 3765 GGAACCTTAAGGAGGCCGCCGATTTAGAGCTTGACGGGGAANAACCGCGACAGCTGGCGA 3824  
|||||  
QY 7729 gaaaggaagggaagaaagcgaagagcggtcgaagcgctggaaggtgtagcgtgca 7788  
|||||  
Db 3825 GAAAGGAGGGAAGAAAGCAAAAGGAGCGGCGCTAGGCGCTGCAAGTGTAGCGGTCA 3884  
|||||  
QY 7789 cgctgacggaaccacacacccgcgcggttatgtgcgcgtacagagcggtgcgcgc 7848  
|||||  
Db 3885 CGCTGGCGTAAACCAACACCGCGCGCTTATGTGCGCTTACAGGCGCGCT---CC 3940  
|||||  
QY 7849 attcgacatcaggtcagcaactgttggaaagcgatcggtgcggcctctcgctat 7908  
|||||  
Db 3941 ATTGCGCATTCAGGCTCGCACTGTGGGAAGGCGATCGGTGCGGCGCTTTCGCTAT 4000  
|||||  
QY 7909 taagcaagctgctcgaaggggggggg 7935  
|||||  
Db 4001 TACGCCAGCTGCGAAAGGGGATGTG 4027  
|||||

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2001, 04:30:13 ; Search time 11937.4 Seconds  
(without alignments)  
3497.675 Million cell updates/sec

Title: US-09-689-430-1\_COPY\_419\_4835

Perfect score: 4417

Sequence: 1 catgaataagagctctcca.....aggcacagagactctactcga 4417

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 10228115 segs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

```
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*
44: gb_est44:*
45: gb_est45:*
46: gb_est46:*
47: gb_est47:*
48: gb_est48:*
49: gb_est49:*
50: gb_est50:*
51: gb_est51:*
52: gb_est52:*
53: gb_est53:*
54: gb_est54:*
55: gb_est55:*
56: gb_est56:*
57: gb_est57:*
58: gb_est58:*
59: gb_est59:*
60: gb_est60:*
61: gb_est61:*
62: gb_est62:*
63: gb_est63:*
64: gb_est64:*
65: gb_est65:*
66: gb_est66:*
67: gb_est67:*
68: gb_est68:*
69: gb_est69:*
70: gb_est70:*
71: gb_est71:*
72: gb_est72:*
73: gb_est73:*
74: gb_est74:*
75: gb_est75:*
76: gb_est76:*
77: gb_est77:*
78: gb_est78:*
79: gb_est79:*
80: gb_est80:*
81: gb_est81:*
82: gb_est82:*
83: gb_est83:*
84: gb_est84:*
85: gb_est85:*
86: gb_est86:*
87: gb_est87:*
88: gb_est88:*
89: gb_est89:*
90: gb_est90:*
91: gb_est91:*
92: gb_est92:*
93: gb_est93:*
94: gb_est94:*
95: gb_est95:*
96: gb_est96:*
97: gb_est97:*
98: gb_est98:*
99: gb_est99:*
100: gb_est100:*
101: gb_est101:*
102: gb_est102:*
103: gb_est103:*
104: gb_est104:*
105: gb_est105:*
106: gb_est106:*
107: gb_est107:*
108: gb_est108:*
109: gb_est109:*
110: gb_est110:*
111: gb_est111:*
112: gb_est112:*
113: gb_est113:*
114: gb_est114:*
115: gb_est115:*
116: gb_est116:*
```



```

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

```

```

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB  | ID       | Description        |
|------------|-------|-------------|--------|-----|----------|--------------------|
| 1          | 467.4 | 10.6        | 573    | 139 | BE755695 | BE755695 209602 MA |
| 2          | 449.2 | 10.2        | 815    | 108 | AU135340 | AU135340 AU135340  |
| 3          | 398.8 | 9.0         | 482    | 138 | BE650938 | BE650938 UT-M-BH3- |
| 4          | 371.2 | 8.4         | 954    | 154 | BC477974 | BC477974 602522966 |
| 5          | 350.8 | 7.9         | 515    | 110 | AV748467 | AV748467 AV748467  |
| 6          | 305.6 | 6.9         | 368    | 136 | BE478981 | BE478981 163493 BA |
| 7          | 263.8 | 6.0         | 517    | 17  | A1176907 | A1176907 EST220513 |
| 8          | 250.2 | 5.7         | 466    | 24  | A1760509 | A1760509 w64604 x  |
| 9          | 249   | 5.6         | 326    | 134 | BE458694 | BE458694 BE458694  |
| 10         | 240   | 5.4         | 308    | 134 | BE449764 | BE449764 BE449764  |
| 11         | 239.8 | 5.4         | 505    | 19  | A1410862 | A1410862 EST239155 |
| 12         | 236.2 | 5.3         | 484    | 148 | BE420450 | BE420450 UT-R-BJ2- |
| 13         | 217.4 | 4.9         | 401    | 141 | BE847128 | BE847128 uw22b11.Y |
| 14         | 216.2 | 4.9         | 314    | 128 | BB218602 | BB218602 BB218602  |
| 15         | 201.4 | 4.6         | 422    | 17  | A1233991 | A1233991 EST230679 |
| 16         | 201.4 | 4.6         | 444    | 116 | AM489638 | AM489638 UT-M-BH3- |
| 17         | 200   | 4.5         | 400    | 22  | A1602045 | A1602045 UT-R-G0-u |
| 18         | 194   | 4.4         | 265    | 3   | AA184901 | AA184901 m153e12.r |
| 19         | 193.4 | 4.4         | 384    | 141 | BE852200 | BE852200 uw22b11.x |
| 20         | 193.2 | 4.4         | 247    | 116 | AM446518 | AM446518 86227 MAR |
| 21         | 192.4 | 4.4         | 313    | 128 | BB219780 | BB219780 BB219780  |
| 22         | 190.2 | 4.3         | 590    | 166 | BE368063 | BE368063 601221956 |
| 23         | 190   | 4.3         | 431    | 116 | AM456831 | AM456831 UT-M-BH3- |
| 24         | 174   | 3.9         | 334    | 116 | BE097510 | BE097510 UT-R-BH1- |
| 25         | 172.2 | 3.9         | 383    | 150 | BE522383 | BE522383 UT-R-C3-S |
| 26         | 158.8 | 3.6         | 927    | 106 | AL552613 | AL552613 AL552613  |
| 27         | 158.4 | 3.6         | 3739   | 192 | AK014835 | AK014835 Mus muscu |
| 28         | 157   | 3.6         | 397    | 226 | AQ270160 | AQ270160 HS_2037-A |
| 29         | 156.2 | 3.5         | 683    | 243 | AZ426543 | AZ426543 IM0208B09 |
| 30         | 152   | 3.4         | 497    | 15  | A1072568 | A1072568 UT-R-C2-n |
| 31         | 150.2 | 3.4         | 286    | 135 | BB475629 | BB475629 BB475629  |
| 32         | 136.6 | 3.1         | 745    | 106 | AL577036 | AL577036 AL577036  |
| 33         | 136.2 | 3.1         | 965    | 146 | BF302800 | BF302800 602032680 |
| 34         | 133.2 | 3.0         | 918    | 106 | AL556703 | AL556703 AL556703  |
| 35         | 131.8 | 3.0         | 864    | 18  | A1324553 | A1324553 m192d06.Y |
| 36         | 131.6 | 3.0         | 807    | 108 | AU134073 | AU134073 AU134073  |
| 37         | 131.2 | 3.0         | 741    | 155 | BE568567 | BE568567 602387557 |
| 38         | 130.6 | 3.0         | 942    | 174 | BG175139 | BG175139 602337457 |
| 39         | 130   | 2.9         | 657    | 17  | A1173222 | A1173222 u663a12.x |
| 40         | 130   | 2.9         | 759    | 154 | BE333260 | BE333260 601860625 |
| 41         | 129   | 2.9         | 1038   | 141 | BE911604 | BE911604 601663080 |
| 42         | 127.8 | 2.9         | 290    | 141 | BE848985 | BE848985 uw01f02.Y |
| 43         | 127   | 2.9         | 857    | 175 | BG242396 | BG242396 602354177 |
| 44         | 125.4 | 2.8         | 608    | 137 | BE554133 | BE554133 ur38608.Y |
| 45         | 124.6 | 2.8         | 775    | 175 | BG246010 | BG246010 602358838 |

## ALIGNMENTS

| RESULT     | 1   | LOCUS       | BE755695 | 573 bp | mRNA | EST | 15-SEP-2000 |
|------------|---|-------------|----------|--------|------|-----|-------------|
| DEFINITION | 209602 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.   |             |          |        |      |     |             |
| ACCESSION  | BE755695  |             |          |        |      |     |             |
| VERSION    | BE755695.1  | GI:10169687 |          |        |      |     |             |
| KEYWORDS   | EST.  |             |          |        |      |     |             |
| SOURCE     | cow.  |             |          |        |      |     |             |
| ORGANISM   | Bos taurus  |             |          |        |      |     |             |
| REFERENCE  | 1 (bases 1 to 573)  |             |          |        |      |     |             |
| AUTHORS    | Smith,T.P.L., Casaa,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Laegreid ,W.M. and Keefe,J.W. |             |          |        |      |     |             |
| TITLE      | Design and use of four pooled tissue normalized cDNA libraries for  |             |          |        |      |     |             |

## JOURNAL COMMENT

EST discovery in cattle

Unpublished (2000)  
Contact: Smith TPLUSDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390Email: smitht@mail.marc.usda.gov  
Single pass sequencing. Bases called and alt\_trimmed with phred v0.980904.e. Vector identified by cross\_match with the -minscore 18 and -minmatch 12 options.

## PCR Primers

FORWARD: AGGAACAGCATGACCAT  
BACKWARD: GTTTTCCAGTCACGAC  
Plate: 59 row: N column: 5  
Seq primer: ATTTAGGACATATAG.  
Location/Qualifiers

## FEATURES

source

1..573

/organism="Bos taurus"

/dp\_xref="taxon:9913"

/clone\_lib="MARC 2BOV"

/tissue\_type="pooled"

/lab\_host="DH10B"

/note="Vector: pCMV SPORT6; Site\_1: XbaI; Site\_2: XhoI; Library made from pooled tissue from testis, thymus, semitendinosus muscle, longissimus muscle, pancreas, adrenal, and endometrium."

BASE COUNT 149 a 134 c 146 g 144 t

## ORIGIN

Query Match 10.6% Score 467.4: DB 139: Length 573:

Best Local Similarity 88.5% Pred. No. 8.3e-123: Matches 507: Conservative 0: Mismatches 66: Indels 0: Gaps 0:

|    |      |   |      |
|----|------|---|------|
| QY | 3089 | ttcactgaaataatgaagaaacacgagcgcctccctcaatccagatgaagatccc           | 3148 |
| DB | 1    | TTTGGCGAAGAAACATGCGAAGACCTGGCGGCGGCTCCATGTCACACGAGAGACCT          | 60   |
| QY | 3149 | acutltaagagaattatcgcttccatgaacaaatgagctacataatgatactacct          | 3208 |
| DB | 61   | ACTTTCAAGAAAGATATCGCTTCATCATCATGCTACGATGATGATGATGATGATGAT         | 120  |
| QY | 3209 | ggcttaagaatggcttgcaggaatgaagatgtatgatactgctcagatgagcaaat          | 3268 |
| DB | 121  | GGCTTAGCATGGCTCAGCATCAAGGATGATGATGATGATGATGATGATGATGATGATGAT      | 180  |
| QY | 3269 | gaacaacatcattatcattatcattatcattatcattatcattatcattatcattatcattat   | 3328 |
| DB | 181  | GAATATATCCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT      | 240  |
| QY | 3329 | tataaatggacatgtacatctatccagtgcttttgaagacagtggaatgtatccaa          | 3388 |
| DB | 241  | TATTAATATGCGGCTGATCAATCTTACCCAGGTCCTTGTGAGAGGTCGAATCTTACCA        | 300  |
| QY | 3389 | tccaaagctgaatttgcgggtggaatgcttttggcgagcatctacatcgtcggagt          | 3448 |
| DB | 301  | TCCAAGGTGGAGCTGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG     | 360  |
| QY | 3449 | agcaacatttcttggttlaagcaataatgagatgagatgagatgagatgagatgagatgagat   | 3508 |
| DB | 361  | AGCACTCTCTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG     | 420  |
| QY | 3509 | caaatagagaatttcaagatcaagcttcaagacataatgagagtgagagtgagagtgagagtg   | 3568 |
| DB | 421  | CGCATTTAAGATTTTTCAGATTTTTCAGATTTTTCAGATTTTTCAGATTTTTCAGATTTTTC    | 480  |
| QY | 3569 | gccagacttaatttccggatcaatgaatgagagtgagagtgagagtgagagtgagagtgagagtg | 3628 |
| DB | 481  | GCCAGACTTATTTTTCGATCAATCAACGCTGAGACCAAGAGATCCCTTCCTTGG            | 540  |
| QY | 3629 | atcaagtgagatctgtgagccaatgattatt                                   | 3661 |

```

Db      541 ATCAAGTGGATCTGTGGCCGATGATTATT 573

RESULT      2
LOCUS       AUI35340
DEFINITION  AUI35340 PLACE1 Homo sapiens CDNA clone PLACE1001786 5', mRNA
sequence.
ACCESSION   AUI35340
VERSION     AUI35340.1
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 815)
AUTHORS     Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isegar,T.
TITLE       HRI human cDNA project
JOURNAL     Unpublished (2000)
COMMENT     Contact: Takao Isegar
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel.: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project: 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES
    source
        1..815
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="PLACE1001786"
        /clone_lib="PLACE1"
        /tissue_type="Placenta"
        /note="Vector: PME18SFL3"
BASE COUNT  200 a 207 c 204 g 197 t 7 others
ORIGIN
Query Match 10.2%; Score 449.2; DB 108; Length 815;
Best Local Similarity 98.3%; Pred. No. 1.7e-117;
Matches 454; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3788 atggtctcttcttggaatggtgattcatctggtgataaaacacaaatattttaacctcca 3847
      |
      |
      |
Db      92 AAGGCTCTTTGGCAATGTCGATTCACTGGGATTAACAAATATTTTAAACCTCCA 151
      |
      |
      |
QY 3848 attattgctgatacatccgcttgcaccaactcattatagatccgcagcactcttcgc 3907
      |
      |
      |
Db      152 ATTTATGCTCATATCATCCGTTTGCAACCACTCATTTATAGATTGCGACGACTCTTCC 211
      |
      |
      |
QY 3908 atggagtgatgaggtgctgatttaaatagttgcagcatgccaattgggaattggaagtaaa 3967
      |
      |
      |
Db      212 ATGGAGTTGATGGCGCTGATTAATAGTTGACAGATGCCATFTGGGAATGAGAGTAA 271
      |
      |
      |
QY 3968 gcaatataagatgcagagattactgcttcacacctacttaacacatatgttgcacctgg 4027
      |
      |
      |
Db      272 GCAATATACAGATGCACAAATTAATCTTCACTTACCAATATGTTTCCACCTGG 331
      |
      |
      |
QY 4028 tctcttcaaaagctcgactcactccaagggaaggaagtaagctgagacctcaagtg 4087
      |
      |
      |
Db      332 TCTCCTTTAAAGCTCGACTTCACCTCCAAAGGAGAGTAATGCTGGAGACCTCAGGTG 391
      |
      |
      |
QY 4088 aataatccaaaagtgctgcaagtgagctccagaagaacaataaagtcacagagta 4147
      |
      |
      |
Db      392 AATATATCAAAAGAGTGCTGCAAGTGGACTTCCAGAAAGACAAATAAAGTCCACAGAGTA 451
      |
      |
      |
QY 4148 actactcagaggaataaactctctgttaaccagcatgtatgtgaagaggttccctcatctcc 4207

```

```

Db      452 ACTACTCAGGAGATAAAATCTGCTTACACAGCATGTATGAAAGACTTCATCCTCC 511
      |
      |
      |
QY 4208 agcagtcacagatgagcatcagtgactcctcttttcagaat 4249
      |
      |
      |
Db      512 AGCAGTCAGATGCCATCATGATGAGCTGCTACCTTCGAATT 553
      |
      |
      |

RESULT      3
LOCUS       BE650938
DEFINITION  BE650938 482 bp mRNA EST 06-SEP-2000
            UI-M-BH3-asm-g-04-0-UI.r1 NIH_BMAP_M.S4 Mus musculus CDNA clone
            UI-M-BH3-asm-g-04-0-UI 5', mRNA sequence.
ACCESSION   BE650938
VERSION     BE650938.1
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 482)
AUTHORS     Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE       Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL     Genome Res. 6 (9), 791-806 (1996)
COMMENT     Contact: Chin, H
            National Institute of Mental Health
            6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
            20892-9643, USA
            Tel: 301 443 1706
            Fax: 301 443 9890
            Email: mestr@mail.nih.gov
            cDNA Library Preparation: M.B. Soares Lab Clone distribution:
            Researchers may obtain BMAP cDNA clones from RESARCH GENETICS. It
            should be noted that Bento Soares is generating a small number of
            additional specialized non-redundant arrays of BMAP cDNAs whose
            availability will be considered under appropriate and limited
            collaborative arrangements
            Seq primer: M13 Reverse.
FEATURES
    source
        1..482
        /organism="Mus musculus"
        /strain="C57BL/6J"
        /db_xref="taxon:10090"
        /clone="UI-M-BH3-asm-g-04-0-UI"
        /clone_lib="NIH_BMAP_M.S4"
        /dev_stage="27-32 days"
        /lab_host="DH10B (Life Technologies)"
        /note="Vector: pT73D-Pac (Pharmacia) with a modified
            polylinker; Site_1: Not I; Site_2: Eco RI; The
            NIH_BMAP_M.S4 library is a subtracted library of a series,
            ultimately derived from a mixture of individually tagged
            normalized libraries from ten regions of the mouse brain
            (cerebellum, brain stems, olfactory bulbs, hypothalamus,
            cortex, amygdala, basal ganglia, pineal gland, striatum,
            hippocampus) after a series of subtractions to reduce the
            representation of cDNAs from which ESTs had already been
            generated. The following serially subtracted libraries
            were generated in this process: NIH_BMAP_M.S4,
            NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, NIH_BMAP_M.S3.1,
            NIH_BMAP_M.S2, NIH_BMAP_M.S1. The subtracted library
            (NIH_BMAP_M.S4) was constructed as follows: PCR amplified
            cDNA inserts from NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, and
            NIH_BMAP_M.S3.1 clones from which 3' ESTs had been derived
            was used as a driver in a hybridization with a pool of
            the NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, and NIH_BMAP_M.S3.1
            libraries in the form of single-stranded circles. The
            remaining single-stranded circles (subtracted library)
            was purified by hydroxyapatite column chromatography,
            converted to double-stranded circles and electroporated
            into DH10B bacteria (Life Technologies) to generate the
            NIH_BMAP_M.S4 library. This procedure has been previously

```

described (Bonaldo, Lennon and Soares, Genome Research  
6:791-806, 1996)"

BASE COUNT 142 a 102 c 101 g 137 t

ORIGIN

Query Match 9.0%; Score 398.8; DB 138; Length 482;  
Best Local Similarity 89.2%; Pred. No. 4.2e-103;  
Matches 430; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 3740 gatggagaagatgagcactatcgaggaattccatggaacctaatggtctctt 3799  
|||||  
Db 1 GATGGGAGAGATGCGCGATTATCAAGGAATCCATCGAACCTTAATGTTCTTT 60  
QY 3800 ggcacatgcatctatgagatgataaacaataatttttaacctcaattatgctga 3859  
|||||  
Db 61 GGCATATGGCATCTATCTGGGATTACCATATATGTTTATCTCCAAATTATGCTCGA 120  
QY 3860 tacatccgtttgacacccaactatataagatcgacagcactctcgcatgagttgatg 3919  
|||||  
Db 121 TATATCGGTTTGACCCCACTCATTTAGCATCGGTAGTACTCTTGCATGAGATTGATG 180  
QY 3920 ggcctgcatatgaatagatgacgacatgcatggaatggaagtaagaacatacagat 3979  
|||||  
Db 181 GCGTGATTTTAAACAGATTCGACGATACCATGGGAATGGAAGTAATATATCAGAT 240  
QY 3980 ggcacagatctactctactactactactactactactactactactactactactact 4039  
|||||  
Db 241 ACACAAATCACTGCTCTATCTATCTTACCAACATGTTCTTACTTGTCTCTTACAA 300  
QY 4040 gctcgacttcaacctcaagaagaggaatgacgctgagagcactgagtgaaataccaaa 4099  
|||||  
Db 301 GCTGACTTACCTCCGAGGAGAGACTAATGCCGCGGACTCAGTGATGATCCAAA 360  
QY 4100 gagtgcgtcgaatgagcttcagagaacaaatgaagtcacaggaataactactagga 4159  
|||||  
Db 361 CAATGCTTGAAGAGGACTTACCAAAAGCAATGAATGCTACGTGAATTAATCCAGGA 420  
QY 4160 gtaaatctctgtctaccagcatgatagtgaagagttctctatctccacagtcagat 4219  
|||||  
Db 421 GTGAAATCTCTTTACCAAGATGTTGTAAGAGATTCCTTATTTTACGACAGTCAAGAT 480  
QY 4220 gg 4221  
||  
Db 481 GG 482

RESULT 4  
LOCUS BG477974 954 bp mRNA EST 21-MAR-2001  
DEFINITION 602522966F1 NIH\_MGC\_20 Homo sapiens cDNA clone IMAGE:4641352 5',  
ACCSSION BG477974.1 GI:13410253  
VERSION BG477974.1 GI:13410253  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 954)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC/DCPD/DPF  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM1408 row: c column: 17

High quality sequence stop: 626.  
FEATURES  
Location/Qualifiers  
1..954  
source

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4641352"  
/clone\_lib="NIH-MGC\_20"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: skin; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adapter: GGCAAGAG(c). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using Zap-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."  
BASE COUNT 233 a 276 c 232 g 212 t  
ORIGIN

Query Match 8.4%; Score 371.2; DB 154; Length 954;  
Best Local Similarity 86.5%; Pred. No. 4.8e-95;  
Matches 585; Conservative 0; Mismatches 68; Indels 23; Gaps 15;

QY 105 atatgcaagtgatctcgtgagctgctgtgagcgaagaattccctcagagtgccaa 164  
|||||  
Db 108 AGATGACCGAGATGCCGCGGAGAAACGTCGCCGGAATTTCTCTAGAGTGCAG 167  
QY 165 aactttccatccaacacctcagctcgtgacaaaagactcgtttgtagaattcacg 224  
|||||  
Db 168 AATCTTTTCATTCAACACTCAGTGTGTACAAAAGACTGTTGTGATTCACGG 227  
QY 225 ttcaaccttcaaatcgtcgaagcgaagcaccctgagatggatctgtagtctacca 284  
|||||  
Db 228 ATCACTTTTCAACATCGCTTAACGCAAGCGCACCTGAGTGGCTGCTAGTCTACCA 287  
QY 285 tccaagctgaggtttttatgatacagtgatcatcaactaagaacatgagcttccatctg 344  
|||||  
Db 288 TCCAGGCTGAGGTTTATGATACAGTGTCTATTAACCTTAAGAACATGCGTCCATCTTG 347  
QY 345 tcaagcttcatctgtgtgtgtatcctactcgtgaagactctga--ggagactgaatga 402  
|||||  
Db 348 TCAGTCTCATGCTGTTGGTGTATCCAGTGAAGCTTGTGACGGAGCCTGAATATGA 407  
QY 403 tgaatgacacagatcaaa--ggagagaagaagatgata-aagttctccgtgtggaagcca 459  
|||||  
Db 408 TGATCAGACCCAGTCAAAACGGGAGACAGAAAGATGATACAGTCTTCCCTGGTGAAGCCA 467  
QY 460 tcaat-atgctctgagagctctg---aaagagaatgltcaaatgagctctgaccactgt 515  
|||||  
Db 468 TACCTCATGCTGTGGCAGGCTCTCAAAAGACATGCTGCCACCTGCTGTGACCACTGA 527  
QY 516 gcc--ttacctactatctcttcatatgagactgtg--taaaagacttgaat--tcagc 571  
|||||  
Db 528 TGCCATTACTACTCATATCTTCTCATGTGACCTGCGTCAAAAGCTTAATCTCAGGA 587  
QY 572 ctcaattggagccttactagatgta--gagaaggagatctgagcaagaanaa--gacacag 628  
|||||  
Db 588 CTCATTGGAGCCCTACTAGTATGATAGAGCAGGAGAGTGTGGCCCAAAACCGACACAG 647  
QY 629 accttgcaca-aatttatctactt--ttgcttatttgaatgaaggaaa--gttggca 684  
|||||  
Db 648 ACCTTGCACACATCTATCTTCTGTGGCTGTATGATGAAGCGAAGAACTTGGCGCA 707  
QY 685 ctcaaga--caagaactcctgatalcag--gataagagatgctgcatctgctggagcctg 741  
|||||  
Db 708 CTGAGAAAGAGAGAACTCTTGTATGACGAGCATAGGATGCTGCTGCTGGGGCCT 767  
QY 742 gctcaaatgacacaga 757  
|||  
Db 768 GGCCTTACATTTGCACA 783

|          |            |   |   |      |     |                |
|----------|------------|---|---|------|-----|----------------|
| RESULT   | 5          |   |   |      |     |                |
| AV748467 | LOCUS      | AV748467  | 515 bp                                  | mRNA | EST | 19-OCT-2000    |
|          | DEFINITION | AV748467  | NPC Homo sapiens cDNA clone NPCAX05 5', |      |     | mRNA sequence. |
|          | ACCESSION  | AV748467  |   |      |     |                |
|          | VERSION    | AV748467.1  | GI:10906315                             |      |     |                |
|          | KEYWORDS   | EST.  |   |      |     |                |
|          | SOURCE     | human.  |   |      |     |                |
|          | ORGANISM   | Homo sapiens  |   |      |     |                |
|          |            | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;   |   |      |     |                |
|          | REFERENCE  | Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  |   |      |     |                |
|          | AUTHORS    | 1 (bases 1 to 515)  |   |      |     |                |
|          |            | Song,H., Peng,Y., Gu,Y., Yang,Y., Gao,G., Xiao,H., Xu,X., Li,N.,  |   |      |     |                |
|          |            | Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., G,   |   |      |     |                |
|          |            | W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Ye,M., Zhang,Q.,   |   |      |     |                |
|          |            | Han,Z., Chen,Z., Hu,R., and Chen,J.   |   |      |     |                |
|          | TITLE      | Homo sapiens NPC library cDNA clones  |   |      |     |                |
|          | JOURNAL    | Unpublished (2000)  |   |      |     |                |
|          | COMMENT    | Contact: Qinghua Zhang<br>Shanghai Institute of Endocrinology, Rui-Jin Hospital<br>197 Rui-jin II Road, Shanghai 200025, P. R. China<br>Tel.: 86-21-66370045(ex.663332)<br>Fax: 86-21-64743206<br>Email: mbshl@sina.com.cn<br>This clone is available at Shanghai Hematology Institute in<br>Shanghai.<br>Chinese National Human Genome Center at Shanghai<br>351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong. |   |      |     |                |

| FEATURES   | SOURCE | Location/Qualifiers  |
|------------|--------|--|
|            |        | 1. .515  |
|            |        | /organism="Homo sapiens"                                       |
|            |        | /db_xref="taxon:9606"  |
|            |        | /clone="NPCAXA05"  |
|            |        | /clone_lib="NPC"   |
|            |        | /tissue_type="pituitary"                                       |
|            |        | /dev_stage="Adult"   |
|            |        | /lab_host="SOLR"   |
|            |        | /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI" |
| BASE COUNT |        | 132 a 114 c 131 g 135 t 3 others                               |
| ORIGIN     |        |  |

| Query Match           | 7.9%  | Score 350.8       | DB 110   | Length 515 |
|-----------------------|---|-------------------|----------|------------|
| Best Local Similarity | 95.3%   | Pred. No. 2.7e-89 |          |            |
| Matches 383           | Conservative 0  | MissMatches 14    | Indels 5 | Gaps 2     |
| QY 3768               | atgcgccttccttggcgaatgttgatctacatcttggaataaacaacatatltttaacctcca | 3847              |          |            |
| Db 111                | AAGGCTCTTCTTGGCAATGTGCATTCATCTGGGATTAACACACATATTTTAACTTC        | 170               |          |            |
| QY 3848               | attatttgcgcgaataccgctttgcaccccaactcattatagaatcgcagcaactcttcgc   | 3907              |          |            |
| Db 171                | ATTATTGGCTCGATACATCCGTTTGCCACCACATCATTTATAGATTGGCAGCACTCTTCCG   | 230               |          |            |
| QY 3908               | atggagcttgatggcgcgtgatctaataagtttcagatatccatttggaaatlgagatlaa   | 3967              |          |            |
| Db 231                | ATGGAGTTGATGGCTGTGATTTAATAATTCTTCAGCATGCCATGGGAATGGAGATTA       | 290               |          |            |
| QY 3968               | gcaaatcagatgcacagaattactgcttcattcttacttaccaatatgtttgcaccttg     | 4027              |          |            |
| Db 291                | GCAATATGAGATGCACAGATTACTGGTTATCTTATCCAAATATGTTGGCACTGG          | 350               |          |            |
| QY 4028               | tctccttccaaagctcgacttcaacctccaaggaggaagtatgctctgagacctcaagt     | 4087              |          |            |
| Db 351                | TCTCCTTCAAAAGCTCAGCTTCACCTTCACAGGAGAGATTAATGCCGAGACCTCAGGTG     | 410               |          |            |
| QY 4088               | aataatccaaaagatggttcgaagtgagagcttccaaagaacatlgaaatgcagagatga    | 4147              |          |            |
| Db 411                | AATTAATCCAAAGAGTGGCTTCAAGTGAAGCTTTCAGAAACATATGCTGACAA-NGAAGTA   | 469               |          |            |
| QY 4148               | actactcagaggaataaactctctgcttaccagcatatgatg                      | 4189              |          |            |

| Db         | 470   | ACTACTTAGAGAGTA---- | CTCTTCTTACACAGCATGTATG         | 507         |
|------------|---|---------------------|--------------------------------|-------------|
| RESULT     | 6   |                     |                                |             |
| BE478981   |   |                     |                                |             |
| LOCUS      |   |                     |                                |             |
| DEFINITION | 163493 BARC   | 5BOV Bos            | taurus CDNA 5', mRNA sequence. | 28-AUG-2000 |
| ACCESSION  | BE478981  |                     |                                |             |
| VERSION    | BE478981.1  | GI:9596514          |                                |             |
| KEYWORDS   | EST.  |                     |                                |             |
| SOURCE     | COW.  |                     |                                |             |
| ORGANISM   | Bos taurus  |                     |                                |             |
|            | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;   |                     |                                |             |
|            | Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;   |                     |                                |             |
|            | Bovidae; Bovinae; Bos.  |                     |                                |             |
|            | 1 (bases 1 to 368)  |                     |                                |             |
| REFERENCE  | Sonstegard,T.S., Capucio,A.V., Van Tassel,J.C.P., Ashwell,M.S. and  |                     |                                |             |
| AUTHORS    | Wells,K.D.  |                     |                                |             |
| TITLE      | Mapping of Expressed Sequence Tags from a normalized bovine mammary |                     |                                |             |
|            | gland cDNA library  |                     |                                |             |
|            | Unpublished (2000)  |                     |                                |             |
| JOURNAL    | Contact: Sonstegard TS  |                     |                                |             |
| COMMENT    | USA, ARS, Beltsville Agricultural Research Center                   |                     |                                |             |
|            | Bldg. 200 Rm 2A, Beltsville, MD 20705, USA                          |                     |                                |             |
|            | Tel: 301 504 8416   |                     |                                |             |
|            | Fax: 301 504 8414   |                     |                                |             |
|            | Email: tadse@psi.barc.usda.gov                                      |                     |                                |             |
|            | Single pass sequencing. Bases called and alt-trimmed with phred     |                     |                                |             |
|            | v0.980904.e. Vector identified by cross_match with the -minscore 18 |                     |                                |             |
|            | and -minmatch 12 options.   |                     |                                |             |
|            | PCR Primers   |                     |                                |             |
|            | FORWARD: AGGAACAGCATGACCAT  |                     |                                |             |
|            | BACKWARD: GTTTCCCACTACGACG  |                     |                                |             |
|            | Plate: 141 row: 6 column: 5   |                     |                                |             |
|            | Seq primer: ATTTAGGTGACACTATAG.                                     |                     |                                |             |

| FEATURES              | SOURCE   | Location/Qualifiers   |
|-----------------------|--|---|
|                       |  | 1..368  |
|                       |  | /organism="Bos taurus"  |
|                       |  | /db_xref="taxon:9913"   |
|                       |  | /clone_id="BARC_5BOV"   |
|                       |  | /tissue_type="pooled"   |
|                       |  | /lab_host="DH10B"   |
|                       |  | /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI; library made from pooled mRNA isolated from mammary tissues at eight physiological, developmental, and disease states." |
| BASE COUNT            | 90 a   | 87 c 98 g 93 t  |
| ORIGIN                |  |   |
| Query Match           | 6.9%;  | Score 305.6; DB 136; Length 368;  |
| Best Local Similarity | 89.4%;   | Pred. No. 2.4e-76;  |
| Matches 329;          | Conservative 0;  | Mismatches 39; Indels 0; Gaps 0;  |
| QY 3329               | tataaatgacacgtgtacaatctctatccagagtgttttttgagacagtgtagaatgttacc   | 3388  |
| Db 1                  | TGTAAATGCGGCGTCTACAACTCTACCCAGAGTGCTTTTGAGACGGTGAATGCTTACCA      | 60  |
| QY 3389               | tccaaagcttggaattggccgggttgaaatgccttatgttcgagacactctacagctcggaatg | 3448  |
| Db 61                 | TCCAAAGTTGGGACTTGGCGGATGAATAAGTGCTTATTGGCGAGCACTVCAAGAGCTGGGATG  | 120   |
| QY 3449               | agcacacttttcctgtgttacaagcaataaggtctcagactcccttcgggaatgcttctga    | 3508  |
| Db 121                | AGCACTCTCTTCCTGCTGGTGTACAGCAGAGAGTGCTCAAAATTCACCTGGGAATGGCTTTGGA | 180   |
| QY 3509               | cacattagagatttcagattacagcttcacagacataatggacagttggccccaagctg      | 3568  |
| Db 181                | CGCATTAAGATTTTCAATATTACACTCTCAGGACATATGGACAGTGGGCCCCCAAGTG       | 240   |
| QY 3569               | gccagactcatlctccgaatcaatcaatgacctgagacacaaggaccccttctctgg        | 3628  |

| Db                    | 241  | GCAGACTTATTATTCGTGATTCATTCACAGCGTGGAGACCAAGATCCCTCTCTTGG      | 300                   |
|-----------------------|--|---|-----------------------|
| OY                    | 3629   | atcaagtgatgactctgttgccaccatgatattcaccgcatcaagaccagggatgccgt   | 3688                  |
| Db                    | 301  | ATCAAGGTGAGATCTGTGGCCGCGATGATTTATTCACAGCATCCTGTACATGAGGGTGC   | 360                   |
| OY                    | 3689   | cagaagtt 3696   |                       |
| Db                    | 361  | CAGAAgTT 368  |                       |
| RESULT                | 7  |   |                       |
| LOCUS                 | A1176907   | 517 bp  | mRNA                  |
| DEFINITION            | EST220513  | Normalized rat ovary, Bento Soares                            | Rattus sp. cDNA clone |
| ACCESSION             | ROVBY42.3  | end, mRNA sequence.   |                       |
| VERSION               | A1176907   |   |                       |
| KEYWORDS              | EST  |   |                       |
| SOURCE                | A1176907.1   | GI:3727545  |                       |
| ORGANISM              | Rattus sp.   |   |                       |
| REFERENCE             | 1 (bases 1 to 517)   |   |                       |
| AUTHORS               | Lee,N.H., Glodex,A., Chandra,I., Mason,T.M., Quackenbush,J., Kerlavage,A.R. and Adams,M.D.   |   |                       |
| TITLE                 | Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat Gene Index   |   |                       |
| JOURNAL               | unpublished (1998)   |   |                       |
| COMMENT               | Contact: Lee, NH<br>The Institute for Genomic Research<br>9712, Medical Center Drive, Rockville, MD 20850, USA<br>Tel: (301)-838-3529<br>Fax: (301)-838-0208<br>Email: nhlee@tigr.org<br>Seq primer: M13-21. |   |                       |
| FEATURES              | Seq primer: M13-21.  |   |                       |
| SOURCE                | Location/Qualifiers  |   |                       |
|                       | 1..517   |   |                       |
|                       | /organism="Rattus sp."   |   |                       |
|                       | /db_xref="ATCC (lnhost):2031594"   |   |                       |
|                       | /db_xref="taxon:10118"   |   |                       |
|                       | /clone="ROVBY42"   |   |                       |
|                       | /clone_id="Normalized rat ovary, Bento Soares"   |   |                       |
|                       | /note="Organ: ovary; Vector: pRT3pac; Site_1: EcoRI; Site_2: NotI"   |   |                       |
| BASE COUNT            | 93 a 174 c 154 g 96 t  |   |                       |
| ORIGIN                |  |   |                       |
| Query Match           | 6.0%; Score 263.8; DB 17; Length 517;  |   |                       |
| Best Local Similarity | 70.9%; Pred. No. 2.9e-64;  |   |                       |
| Matches               | 365; Conservative 0; Mismatches 147; Indels 3; Gaps 1  |   |                       |
| OY                    | 2474   | cagagcccccagcttccaaagaacacagacactattatgtctgagtgagag           | 2533                  |
| Db                    | 513  | CAGGACCCCGGGCTTTCAGCAGAAACGGCGCACTATTTCATGCGCCCGTGAAGGAG      | 454                   |
| OY                    | 2534   | cctctggaattaggaatagtagtgcctcccaactgtttcctaagaagaaggctcaagtgtc | 2593                  |
| Db                    | 453  | CTTTGGGACTACGGGGGTGCGTGTGCGCTGAGCTCGGGGACAGGGCGTGGAGTGGG      | 394                   |
| OY                    | 2594   | agtgccctcagttcaagaagttgtttccagagaatttactgatatgctcccttactcag   | 2653                  |
| Db                    | 393  | GACGACGCTAGGTTCGGGAAAGTGTTCCTGATTCACGAGGGGTCTTCACGCGAG        | 334                   |
| OY                    | 2654   | ccctataaccgttggaacataatgaaacatttgggactctcggggccatatataagagca  | 2713                  |
| Db                    | 333  | CGCGGACACCGGGGAGCTGGACACACACACCTCGGGGCTACTGGGGCCATACATACGGCG  | 274                   |
| OY                    | 2714   | gaagttgaagaataatatatgttacttcagaataaataagagcctctcgtccctattcttc | 2773                  |
| Db                    | 273  | GAGGTGAGGACACATGTGTGTGACTTCCGGAAACAGAGCTCCCGCCCTACTACTCTTC    | 214                   |

|                       |  |   |                       |
|-----------------------|--|---|-----------------------|
| QY                    | 2774   | tattcttagccttatcttctatgaaggaataatcagagcgcaagagcgacgaagcattagaanaaac | 2833                  |
| Db                    | 213  | TACTCCAGCCTCGTGTCTCTACCCGGAGCCGACGCT---                             | GGAGCTGCACCAAGAGCAAC  |
| QY                    | 2834   | ttgttcgaagccttaagaaacccaacttacttlttgaaagtcaacatcatatgtgacc          | 2893                  |
| Db                    | 156  | TTTTTGGGGGCCCAAGAGCAACCAAACTTACTTCTGTGAGAGTCCGGCCCAACATGTGGCC       | 97                    |
| QY                    | 2894   | actaaagatagtttgaactgtcaagcctgggcttattctctgtatgttgaacttgtaaaa        | 2953                  |
| Db                    | 96   | ACGGACGGGAGATTGTGACTGTGCAGAGCCTGGGGCTATTCTTCGACGTGACTCGAGCGA        | 37                    |
| QY                    | 2954   | gatgtcaactgaagcctatitgacccttcggt                                    | 2988                  |
| Db                    | 36   | GACCTGCACCTCGGGCTGATTGGGCTACTTCTGCT                                 | 2                     |
| RESULT                | 8  |   |                       |
| LOCUS                 | A1760509/c   |   |                       |
| DEFINITION            | A1760509   | 466 bp  | mRNA                  |
| ACCESSION             | wh4d04.x1  | NCI_CGAP_Kid11 Homo sapiens   | CDNA IMAGE:2385511 3' |
| VERSION               | A1760509   | similar to g9:MT4113 COAGULANTIN FACTOR VIII PRECURSOR (HUMAN).,    |                       |
| KEYWORDS              | A1760509   | mRNA sequence.  |                       |
| SOURCE                | A1760509.1   | GI:5176176  |                       |
| ORGANISM              | human.   |   |                       |
| REFERENCE             | Homo sapiens   |   |                       |
| AUTHORS               | Eukariyola, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |   |                       |
| TITLE                 | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.         |   |                       |
| JOURNAL               | NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.                      |   |                       |
| COMMENT               | National Cancer Institute, Cancer Genome Anatomy Project (CGAP),   |   |                       |
|                       | Tumor Gene Index   |   |                       |
|                       | Unpublished (1997)   |   |                       |
|                       | Contact: Robert Strausberg, Ph.D.                                  |   |                       |
|                       | Email: c9qab5-remail.nih.gov                                       |   |                       |
|                       | Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  |   |                       |
|                       | Emmert-Buck, M.D., Ph.D.   |   |                       |
|                       | CDNA Library Preparation: M. Bento Soares, Ph.D.                   |   |                       |
|                       | CDNA Library Arrayed by: Greg Lennon, Ph.D.                        |   |                       |
|                       | DNA Sequencing by: Washington University Genome Sequencing Center  |   |                       |
|                       | Clone distribution: NCI-CGAP clone distribution information can be |   |                       |
|                       | found through the I.M.A.G.E. Consortium/LNLN at:                   |   |                       |
|                       | www.bio.lnl.gov/bbrp/image/image.html                              |   |                       |
|                       | Insert Length: 1564  | Std Error: 0.00   |                       |
|                       | Seq primer: -400p from Glbco                                       |   |                       |
|                       | High quality sequence stop: 404.                                   |   |                       |
| FEATURES              | Location/Qualifiers  |   |                       |
| source                | 1..466   |   |                       |
|                       | /organism="Homo sapiens"   |   |                       |
|                       | /db_xref="taxon:9606"  |   |                       |
|                       | /clone="IMAGE:2385511"   |   |                       |
|                       | /clone_id="NCI_CGAP_Kid11"   |   |                       |
|                       | /lab_host="DH10B"  |   |                       |
|                       | /note="Organ: kidney; Vector: pUT7n3D-Pac (Pharmacia) with         |   |                       |
|                       | a modified polylinker; Site:1: Not I; Site:2: Eco RI;              |   |                       |
|                       | Plasmid DNA from the normalized library NCI_CGAP_Kid3              |   |                       |
|                       | was prepared, and ss circles were made in vitro. Following HAP     |   |                       |
|                       | purification, this DNA was used as tracer in a subtractive         |   |                       |
|                       | hybridization reaction. The driver was PCR-amplified cDNAs         |   |                       |
|                       | from a pool of 5,000 clones made from the same library             |   |                       |
|                       | (cloneids 13223376-1323911, 1456007-1456775, and                   |   |                       |
|                       | 1500552-1502855)."   |   |                       |
|                       | Subtraction by Bento Soares and M.                                 |   |                       |
|                       | Fatima Bonaldo.  |   |                       |
| BASE COUNT            | 128 a  | 96 c  | 77 g                  |
| ORIGIN                |  |   | 165 t                 |
| Query Match           | 5.7%   | Score 250.2;  | DB:24; Length 466;    |
| Best Local Similarity | 95.2%  | Pred. No. 2.3e-60;  |                       |
| Matches 258;          | Conservative   | 0;  | Mismatches 13;        |
|                       |  | Indels  | 0;                    |
|                       |  | Gaps  | 0;                    |

| QY         | 1197  | aaacttggtgaactaaccttgctgctgaagaagagagactggagactatgctcccttaagctc | 1256 |
|------------|---|---|------|
| Db         | 453   | AAACTTGGGTACATTACATTCTGCTGCTGAGAGAGAGAGACTGGACTATGCTCCCTTAATGCC | 394  |
| QY         | 1257  | tcgcccccgatgcagagaagtltataaagtcacatcttgacacaaatggccccccagcgatg  | 1316 |
| Db         | 393   | TGCGCCCGGATGACAGAAAGTTATTAAGAAGTCATTATTGAAACAATGGCCCTCAGCGGATGG | 334  |
| QY         | 1317  | gtagaagaattacaaaaaagtcgcgattatggcctacacagatgaataaccttaagactcgtg | 1376 |
| Db         | 333   | GTAGAAGATGACAAAAAAGTCGATTTATGCGCTACACACATGATTAACCTTTAAGACTCGCG  | 274  |
| QY         | 1377  | aagcctatcagatgaatcagaagactctggagcccttacttatgagggaagtctggagaca   | 1436 |
| Db         | 273   | AAGCTATTTCAGCATGAAATCAGAAATCTTGGGACCTTTACTTTATGCGGAAGTTGGAGACA  | 214  |
| QY         | 1437  | cactgttgattatatttaagaatcaagcaag                                 | 1467 |
| Db         | 213   | CACGTGGTAGTTGAGAGAAAAAGATTTTAG                                  | 183  |
| RESULT     | 9   |   |      |
| BB458694   |   |   |      |
| LOCUS      | 326 bp  | mRNA  | EST  |
| DEFINITION | BB458694 RIKEN full-length enriched, 12 days embryo spinal ganglion   |   |      |
| VERSION    | Mus musculus cDNA clone D13008B20.3, similar to I05513 Mus  |   |      |
| KEYWORDS   | domesticus coagulation factor VIII mRNA, mRNA sequence.   |   |      |
| ACCESSION  | BB458694  |   |      |
| VERSION    | BB458694.1  | GI:9354187  |      |
| KEYWORDS   | EST.  |   |      |
| SOURCE     | house mouse.  |   |      |
| ORGANISM   | Mus musculus  |   |      |
| REFERENCE  | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. I (bases 1 to 326)   |   |      |
| AUTHORS    | Kono, H., Aikawa, K., Akahira, S., Akiyama, J., Arahawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomimaga, N., Toya, T., Tsunoda, Y., Watabiki, A., Watanabe, S., Yamamura, T., Yamana, I., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  |   |      |
| TITLE      | RIKEN Mouse ESTs (Kono, H., et al.)   |   |      |
| JOURNAL    | Unpublished (2000)  |   |      |
| COMMENT    | Contact: Yoshihide Hayashizaki<br>Genome Exploration Research Group, Life Science Tsukuba Center,<br>The Institute of Physical and Chemical Research (RIKEN), Genomic<br>Sciences Center<br>3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan<br>Tel: +81-298-36-9013<br>Fax: +81-298-36-9098<br>Email: genome-res@rc.riken.go.jp/<br>URL: http://genome.riken.go.jp/<br>Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasakihara, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.<br>Thermotranscription and thermoinactivation of thermolabile enzymes by<br>trichostatin and its application for the synthesis of full length<br>cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)<br>Itoh, M., Kikuchi, N., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,<br>Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki,<br>Y., and Hayashizaki, Y.<br>Automated filtration-based high-throughput plasmid preparation<br>system. Genome Res. 9 (5), 463-470 (1999)<br>Carninci, P. and Hayashizaki, Y.<br>High-efficiency full-length cDNA cloning. Methods Enzymol. 303,<br>19-44 (1999) |   |      |

Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

Location/Qualifiers  
1..326  
source

```

/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="D130058E20"
/clone_1lb="RIKEN full-length enriched, 12 days embryo spinal ganglion"
/tissue_type="spinal ganglion"
/dev_stage="12 days embryo"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGACGAGAGCAGCGCCGCCAACCTGCAGCTTTTTTTTTTTTVVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAAGAGATTCGTGAGATTAAATTAATTAATTCCTCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pbluescript KS(+) after bulk excision from Lambda FLX I."

```

|                       |              |                    |                |             |
|-----------------------|--------------|--------------------|----------------|-------------|
| Query Match           | 5.6%         | Score 249;         | DB 134,        | Length 326; |
| Best Local Similarity | 86.0%;       | Pred. No. 4.3e-60; |                |             |
| Matches 276;          | Conservative | 0;                 | Mismatches 45; | Indels 0;   |
|                       |              |                    |                | Gaps 0;     |

Oy 740 tggcctaataatgcacacagtcattgttatgtlaaacagagctctcgcagctctgatyyga 799  
 || ||| || ||||| ||||| || | ||||| ||| || ||||| ||||| ||||| |||||  
 Db 2 TGCCCTTTTATCCACACTTTCAAAGGCATCTTAAACAGCTCTTTTCCAGGCTGATTTGA 61

Oy 800 tgcacaggaatcagtctattgycatgtgatggaatgygcaccactctgaatgcaac 859  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
db 62 TACCAATAGGAATACGTCTACTGGCACCCTGATTCGATGGCTCCACCTCGAATACAC 121

860 tcataatcctcgaagtcacacattcttgytgaagaacatcgccaggctctcggaa 919

920 atctcgccaataacttctcctaactgtctcaaacacatcttgatgaccttggaagttcta 979  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

980 ctgttttgcatatctcttccccaacatgagcgatggaagcttatgtcaagtagac 1033

Db 242 CATTATTTTCATTCTCTCCCATAAACATGATGGCAATGGAAGCTTATGTCAAAGTAGAT 301

Oy 1040 agctgtccagagaaacccaa 1060

Db 302 AGCTGCCCTGAGGAATCCCA 322

| LOCUS      | 308 bp   | mRNA | EST | 21-JUL-2000 |
|------------|--|------|-----|-------------|
| DEFINITION | RIKEN full-length enriched, 9 days embryo Mus musculus |      |     |             |
| BB449764   |  |      |     |             |
| BB449764   |  |      |     |             |

|            |  |
|------------|--|
| ACCESSION  | U05576   |
| VERSION    | 1  |
| DEFINITION | CDNA clone D0300/3H24.3, similar to U05573 Mus domesticus coagulation factor VIII mRNA, mRNA sequence. |
| KEYWORDS   | BB449764   |
| FEATURES   | BB449764.1. C1-0310700   |

|          |              |
|----------|--------------|
| KEYWORDS | EST.         |
| SOURCE   | house mouse. |
| ORGANISM | Mus musculus |

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;



1 (bases 1 to 308)  
Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci

| TITLE  | JOURNAL                                    | COMMENT  |
|--|--|--|
| <p>P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Iizawa, M., Kadori, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shiimoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomihata, N., Toyota, T., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yamana, I., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, T.</p> | <p>RIKEN Mouse ESTs (Kono, H., et al.)</p> | <p>Unpublished (2000)<br/>Contact: yoshinide Hayashizaki</p> |

RIKEN Mouse ESTs (Konno, H., et al.)  
Unpublished (2000)

Contact: Yoshihide Hayashizaki  
Genome Exploration Research Group, Life Science Tsukuba Center,  
Genome Science Laboratory  
The Institute of Physical and Chemical Research (RIKEN), Genomic  
Sciences Center  
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
Tel: +81-298-36-9013  
Fax: +81-298-36-9098  
Email: genome-res@rtc.riken.go.jp,  
url: <http://genome.rtc.riken.go.jp/>,  
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki,  
N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
Thermotranscription and thermotranscription of thermolabile enzymes by  
thermolase and its application for the synthesis of full length  
cDNA. *Proc. Natl. Acad. Sci. U.S.A.* 95 (2), 520-524 (1998)  
Itoh, M., Katsunuma, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,  
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki,  
Y. and Hayashizaki, Y.  
Automated filtration-based high-throughput plasmid preparation  
system. *Genome Res.* 9 (5), 463-470 (1999)  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning. *Methods Enzymol.* 303,  
19-44 (1999)  
Please visit our web site (<http://genome.rtc.riken.go.jp>) for  
further details.

|        |                     |
|--------|---------------------|
| JRES   | Location/Qualifiers |
| source | 1. .308             |

```

BASE COUNT      85 a      80 c      50 g      93 t
ORIGIN
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="D030073H24"
/clone_lib="RIKEN full-length enriched, 9 days embryo"
/dec_stage="9 days embryo"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Science Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAAGAGAGAGCGCGCCGCAACTCGAGTCTTTTCTTTTCTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAAGAGAGATCTCTCGACTTATTTAATTAATTCGCCCCCCCCCC 3']. cDNA
was cleaved with BamHI and XhoI. Vector: a modified
pBluescript KS(+) after bulk excision from Lambda FLC I."

```

|      |      |      |      |
|------|------|------|------|
| 85 d | 80 c | 50 g | 93 t |
|------|------|------|------|

```
Query Match      5.4%; Score 240; DB 134; Length 308;
Best Local Similarity 86.2%; Pred. No. 1.6e-57;
Matches 261; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

|    |      |   |      |
|----|------|---|------|
| Db | 9    | GCATGTAAACAGTCTCTTCACGTTCTGATTGGATCCCATGGAATACAGTCTACTGGC         | 68   |
| Qy | 825  | atgtgatgttgaaatggygcacaccccttcgaagtgcactcaatatctctgaaagttacacat   | 884  |
| Db | 69   | ACCGATGTGGAATGGGGACACCTCCGAAATACACCAATATTTCTCGAAGGTCCACAT         | 128  |
| Qy | 885  | tctctgtgaaagaacatcgcacagcgctctcttgaaatcgcgcaataacttctctactg       | 944  |
| Db | 129  | TTTTTGTGAGGAGAACCAOCCCTCAAGCTTCATTGTGAGATATCCACCAATTACTTCCCTTACTG | 188  |
| Qy | 945  | ctcaaacactcttgatgagaccttgaagatcttctactglttgtcatalactcttccacc      | 1004 |
| Db | 189  | CTCAAAACACTCTTGAGATGATCTTGGCCACTCTCTACTATTTTTCATTCTCTCTCCCATTA    | 248  |
| Qy | 1005 | aacactgtgcatggaagactatgtccaagtgtgaacagctgctccagagaacccca          | 1060 |
| Db | 249  | AACCTAGTGGCATGGAAGGCTTATTTCAAAGTAGATATGCTCCCTCTAGGAGATGCCAA       | 304  |

|            |  |                     |      |     |
|------------|--|---------------------|------|-----|
| LOCUS      | A1410862   | 505 bp              | mRNA | EST |
| DEFINITION | EST229155 Normalized rat heart. Bento Soares Rattus sp. cDNA clone |                     |      |     |
| ACCESSION  | RHEC209.3'   | end, mRNA sequence. |      |     |
| VERSION    | A1410862   |                     |      |     |
| KEYWORDS   | A1410862.1   | GI:4254366          |      |     |
| SOURCE     | Rattus sp.   |                     |      |     |

Rattus sp.  
Rattus sp.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 505)  
Lee, N.H., Glodek, A., Chandra, I., Mason, T.M., Quackenbush, J.,

| TITLE  | Gene Index |
|--|------------|
| Kerlavage, A.R. and Adams, M.D.                                  |            |
| Rat Genome Project: Generation of a Rat EST (RESI) Catalog & Rat |            |

Unpublished (1998)  
Contact: Lee, NH

The Institute for Genomic Research  
9712, Medical Center Drive, Rockville, MD 20850, USA

Location/Qualifiers  
1. .505

Email: [nhlee@tigr.org](mailto:nhlee@tigr.org)  
Seq primer: M13-21.

```
1. .505
/organism="Rattus sp."
```

```

      /db_xref="taxon:10118"
      /clone="RHEC209"
      /clone_l1p="Normalized rat heart, Bento Soares"
      /note="Organ: heart; Vector: pTf73pac; Site_1: EcoRI
      Site_2: NotI"
BASE COUNT      83 a      148 c      165 g      109 t
ORIGIN

```

|      |       |       |       |
|------|-------|-------|-------|
| 83 a | 148 c | 165 g | 109 t |
|------|-------|-------|-------|

|                           |       |                   |          |            |
|---------------------------|-------|-------------------|----------|------------|
| Query Match               | 5.4%  | Score 239.8       | DB 19    | Length 505 |
| Best Local Similarity     | 69.6% | Pred. No. 2.3e-57 |          |            |
| Matches 357; Conservative | 0     | Mismatches 147    | Indels 9 | Gaps 2     |

[illegible]

subtraction has been previously described in (Bonaldto, Lennon and Soares, Genome Research 6:791-806, 1996)

TAG\_SEQ=None found"

|            |      |       |       |      |
|------------|------|-------|-------|------|
| BASE COUNT | 98 a | 137 c | 164 g | 85 t |
| ORIGIN     |      |       |       |      |

Query Match 5.3%; Score 236.2; DB 148; Length 484;  
Best Local Similarity 68.2%; Pred. No. 2.4e-56;  
Matches 328; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

|    |      |   |      |
|----|------|---|------|
| Qy | 1253 | gtccctgcccccgatgaagaagtattataaagtcacatatttgacaatgacctgaagcgg      | 1312 |
| Db | 4    | gccccacccttgaagagacggagccctggcggaacccgtaacctggggcgccccctggccgc    | 63   |
| Qy | 1313 | attgttagaagatcacaaaaaagtcgatttatgacatatacagatgaaaccttttagact      | 1372 |
| Db | 64   | atccgcacacaaatattaaagaaagtcaggttcgtgcgtcacacgagcggcaccctttcacacg  | 123  |
| Qy | 1373 | cgtagaactatcagcatgaatcagaatccttggagaccttaacttatalggggaattga       | 1432 |
| Db | 124  | cgtgagccaccaccccgctgaacctggcgctctttagggccactgcctttacggggagctcgga  | 183  |
| Qy | 1433 | gacacactgttgatataattaagaatcaagaacagacacatataacatctaccctcac        | 1492 |
| Db | 184  | gactcactggttgatgtgtttcaagaaacgagcagcaccggcgctatcaacatccaccctcat   | 243  |
| Qy | 1493 | ggaatcactgattgtccgtcctttgtattcagaagaattaccaaagtggtaaacatttg       | 1552 |
| Db | 244  | gggattccctgacgtttggcgctgtgcattgcaggagagcctgccacgaggggtgaagcacgttg | 303  |
| Qy | 1553 | aaggaatttcccaattcttcagagagaataatcaataataatgacagtgactgtagaa        | 1612 |
| Db | 304  | aaggaatttgcgcatccggccacgagaaagacgttcaagtgtaagctgacacctgacccagag   | 363  |
| Qy | 1613 | gatggccaactaaatcagatcctctggtgcttgaccgctattacttagtttgtaat          | 1672 |
| Db | 364  | gatggcccttgcaaggttcggaatgcgggggtgtgagacccttactatgacagcccttgatg    | 423  |
| Qy | 1673 | atggagagagatctagcttcagactcaattggccctccctccatctgctacaaagaatct      | 1732 |
| Db | 424  | cccgagagagaccttgcttccgggctgattggcccccttctatcttgctgtaaaaatct       | 483  |
| Qy | 1733 | g   | 1733 |
| Db | 484  | g   | 484  |

|            |  |
|------------|--|
| RESULT     | 13   |
| LOCUS      | BE847128   |
| DEFINITION | BE847128 401 bp mRNA   |
| LOCATION   | uw22b1.1 Soares mouse 3NDMS Mus musculus cDNA clone IMAGE:3417405  |
| ACCESSION  | 5' similar to gb:U0573 Mus domesticus coagulation factor VIII mRNA |
| VERSION    | BB847128   |
| KEYWORDS   | BB847128.1 GI:10305395   |
|            | EST.   |

**SOURCE** house mouse.  
**ORGANISM** Mus musculus.  
**REFERENCE** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 401)  
**AUTHORS** NCI-CCAG <http://www.ncbi.nlm.nih.gov/ncicgag>.  
**TITLE** National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
**JOURNAL** Unpublished (1997)  
**COMMENT** Contact: Robert Strausberg, Ph.D.  
Email: [cga@bms.fda.gov](mailto:cga@bms.fda.gov)  
This clone is available royalty-free through LINL ; contact the IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
MGI:1093217  
Seq primer: -40RP from Gibco.  
Location/Qualifiers



|    |      |  |      |
|----|------|--|------|
| Oy | 2150 | CGAACCAAGGCATCACCGCCTCACTAAGTTCTTCAGTTCTGTAAACAAGACACGCTGAT    | 2209 |
| Db | 121  | CGAAGACAGAGGTATGACACGATTCGCTGAAAGTTCTTGTTGTGACTGAGGCACCTAGTGAT | 180  |
| Oy | 2210 | TATCAAGAGACAGTCTATCAAGATATTTCCAGCATACCTGTCGAGTAAACAATGCGCAT    | 2269 |
| Db | 181  | TATTTATGAAGAAATTATATGAAGATTTTCCACACAGTTGGTGATGAGAACCAATGTCATT  | 240  |
| Oy | 2270 | GAACCAAGAAGCTCTCCAGAAATCAAGACACCGCTGACCTAGGCAAAAGCAATTAAT      | 2329 |
| Db | 241  | GATCCCAAGAGCTTCTTCCAAATACAAATATCATCTATTACTAGGAAAAAGAAATTCAAA   | 300  |
| Oy | 2330 | GCCACCCCA  | 2338 |
| Db | 301  | GATTCACCA  | 309  |

| RESULT     | 15   |
|------------|--|
| LOCUS      | AI233991   |
| DEFINITION | AI233991 422 bp mRNA EST230679 Normalized rat lung, Bento Soares Rattus sp. cDNA clone |
| ACCESSION  | AI233991   |
| VERSION    | AI233991.1   |
| KEYWORDS   | EST.   |
| SOURCE     | Rattus sp.   |
| ORGANISM   | Rattus sp.   |

REFERENCE 1 (bases 1 to 422)  
AUTHORS Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,  
Kerlavage,A.R. and Adams,M.D.  
TITLE Rat Genome Project: Generation of a Rat EST (REEST) Catalog & Rat  
Gene Index  
JOURNAL Unpublished (1998)  
COMMENT On Oct 30, 1998 this sequence version replaced g1:3817871.

JOURNAL Unpublished (1998)  
 COMMENT On Oct 30, 1998 this sequence version replaced gi:3817871.  
 Contact: Lee, NH  
 The Institute for Genomic Research  
 9712, Medical Center Drive, Rockville, MD 20850, USA  
 Tel: (301)-838-3529  
 Fax: (301)-838-0208  
 Email: nhlee@igf.org  
 Seq primer: M13-21.

| FEATURES | Location/Qualifiers |
|----------|---------------------|
| source   | 1. .422             |

```
/organism="Rattus sp."  
/db_xref="taxon:10118"
```

```

/clone="RLUCT80"
/clone_lib="Normalized rat lung, Bento Soares"
/notes="Organ: Lung; Vector: pT73Pac; Site_1: EcoRI
Site_2: NotI"

```

|            |      |       |       |      |
|------------|------|-------|-------|------|
| BASE COUNT | 74 a | 120 c | 160 g | 68 t |
| ORIGIN     |      |       |       |      |

|             |       |              |        |             |
|-------------|-------|--------------|--------|-------------|
| Query Match | 4.68; | Score 201.4; | DB 17; | Length 422; |
|-------------|-------|--------------|--------|-------------|

Best Local Similarity 67.5%; Pred. No. 2.3e-46;  
Matches 283; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

```

Qy 3033 cagtaacaggaatttgtctctgttlltcaacalcatttgatagaccaaagctgtacttca 3092
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3 CAGTCAAGGAGTTCCGCGCTCTTTCGGCGTGTTCGATGACACCAAGAGCTGTACTTCG 62

```

```

Oy  3093 ctgaataatgtaaaagaacctgcacgctccctgaataatccagatgtaaatccactt 3152
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  63  CTGAGAACCTGACACCGAGTGCAGCCGCCCGCGGACACCCCGCGGAGACCCCGCCT 122

```

Qy 3153 ttaagagaattatcgcttcacatgcaatgactacataatgatactactcgtgc 3212

Db 123 GGAGAGAGAGATCCGCTTCATGCGGTGAATGGTTACGTCGCACATGCTTGGCC 182

3213 tagtaatggtcagatcaagatcgtatctgctcagcatgtggcagcaatgaaa 3272

Db 183 TGGTGAATGCGCGAAGGTGGAGGACGAGGTGGCATCTGCTCAACATGGGCGACCCCGGGC 242

Qy 3273 aacatcattcattcattcagtcgagacatgltgttcacgttgcagcaaaaagaagcagata 3333

Db 243 ACGCACTCCGTGCATTTCAGCGCGACACAGCGTACCGTCGGCGGACGGGGCGAGACACA 302

Qy 3333 aatgtgcactgtacaatctcattccatccagtgltttttgagacagtgcgaatgttaccatcca 3392

Db 303 GGACGGCGGTGTGCAACCTGTACCCAGCGGTGTTACACACCGTGGAGATGTTGCCGTCA 362

Qy 3393 aagtcgaaatttgcgcggttgcgaatgccttaattgcgcagcatctacatgtcttgatgagc 3451

Db 353 GAGCGGGAATTGGCGGGGTGGAGAGCCTGCTGGGCGAGACACTACCGGCTGGGATGAGC 421

Search completed: November 18, 2001, 04:30:38  
Job time: 39920 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2001, 04:30:38 ; Search time 11937.4 Seconds  
(without alignments)  
3773.245 Million cell updates/sec

Title: US-09-689-430-1\_COPY\_150\_4914

Perfect score: 4765

Sequence: 1 cctcttcaagtaaacagta.....ttgggtcgttgcgcgac 4765

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 10228115 segs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.\*

```
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*
44: gb_est44:*
45: gb_est45:*
46: gb_est46:*
47: gb_est47:*
48: gb_est48:*
49: gb_est49:*
50: gb_est50:*
51: gb_est51:*
52: gb_est52:*
53: gb_est53:*
54: gb_est54:*
55: gb_est55:*
56: gb_est56:*
57: gb_est57:*
58: gb_est58:*
59: gb_est59:*
60: gb_est60:*
61: gb_est61:*
62: gb_est62:*
63: gb_est63:*
64: gb_est64:*
65: gb_est65:*
66: gb_est66:*
67: gb_est67:*
68: gb_est68:*
69: gb_est69:*
70: gb_est70:*
71: gb_est71:*
72: gb_est72:*
73: gb_est73:*
74: gb_est74:*
75: gb_est75:*
76: gb_est76:*
77: gb_est77:*
78: gb_est78:*
79: gb_est79:*
80: gb_est80:*
81: gb_est81:*
82: gb_est82:*
83: gb_est83:*
84: gb_est84:*
85: gb_est85:*
86: gb_est86:*
87: gb_est87:*
88: gb_est88:*
89: gb_est89:*
90: gb_est90:*
91: gb_est91:*
92: gb_est92:*
93: gb_est93:*
94: gb_est94:*
95: gb_est95:*
96: gb_est96:*
97: gb_est97:*
98: gb_est98:*
99: gb_est99:*
100: gb_est100:*
101: gb_est101:*
102: gb_est102:*
103: gb_est103:*
104: gb_est104:*
105: gb_est105:*
106: gb_est106:*
107: gb_est107:*
108: gb_est108:*
109: gb_est109:*
110: gb_est110:*
111: gb_est111:*
112: gb_est112:*
113: gb_est113:*
114: gb_est114:*
115: gb_est115:*
116: gb_est116:*
```

117: gb\_est48:\*  
118: gb\_est49:\*  
119: gb\_est50:\*  
120: gb\_est51:\*  
121: gb\_est52:\*  
122: gb\_est53:\*  
123: gb\_est54:\*  
124: gb\_est55:\*  
125: gb\_est56:\*  
126: gb\_est57:\*  
127: gb\_est58:\*  
128: gb\_est59:\*  
129: gb\_est60:\*  
130: gb\_est61:\*  
131: gb\_est62:\*  
132: gb\_est63:\*  
133: gb\_est64:\*  
134: gb\_est65:\*  
135: gb\_est66:\*  
136: gb\_est67:\*  
137: gb\_est68:\*  
138: gb\_est69:\*  
139: gb\_est70:\*  
140: gb\_est71:\*  
141: gb\_est72:\*  
142: gb\_est73:\*  
143: gb\_est74:\*  
144: gb\_est75:\*  
145: gb\_est76:\*  
146: gb\_est77:\*  
147: gb\_est78:\*  
148: gb\_est79:\*  
149: gb\_est80:\*  
150: gb\_est81:\*  
151: gb\_est82:\*  
152: gb\_est83:\*  
153: gb\_est84:\*  
154: gb\_est85:\*  
155: gb\_est86:\*  
156: gb\_est87:\*  
157: gb\_est88:\*  
158: gb\_est89:\*  
159: gb\_est90:\*  
160: gb\_est91:\*  
161: gb\_est92:\*  
162: gb\_est93:\*  
163: gb\_est94:\*  
164: gb\_est95:\*  
165: gb\_est96:\*  
166: gb\_est97:\*  
167: gb\_est98:\*  
168: gb\_est99:\*  
169: gb\_est100:\*  
170: gb\_est101:\*  
171: gb\_est102:\*  
172: gb\_est103:\*  
173: gb\_est104:\*  
174: gb\_est105:\*  
175: gb\_est106:\*  
176: gb\_est107:\*  
177: gb\_est108:\*  
178: gb\_est109:\*  
179: gb\_est110:\*  
180: gb\_est111:\*  
181: gb\_est112:\*  
182: gb\_est113:\*  
183: gb\_est114:\*  
184: gb\_est115:\*  
185: gb\_est116:\*  
186: gb\_est117:\*  
187: gb\_est118:\*  
188: gb\_est119:\*  
189: gb\_est120:\*

190: gb\_est121:\*  
191: gb\_est122:\*  
192: gb\_est123:\*  
193: gb\_est124:\*  
194: gb\_est125:\*  
195: gb\_est126:\*  
196: gb\_est127:\*  
197: gb\_est128:\*  
198: gb\_est129:\*  
199: gb\_est130:\*  
200: gb\_est131:\*  
201: gb\_est132:\*  
202: gb\_est133:\*  
203: gb\_est134:\*  
204: gb\_est135:\*  
205: gb\_est136:\*  
206: gb\_est137:\*  
207: gb\_est138:\*  
208: gb\_est139:\*  
209: gb\_est140:\*  
210: gb\_est141:\*  
211: gb\_est142:\*  
212: gb\_est143:\*  
213: gb\_est144:\*  
214: gb\_est145:\*  
215: gb\_est146:\*  
216: gb\_est147:\*  
217: gb\_est148:\*  
218: gb\_est149:\*  
219: gb\_est150:\*  
220: gb\_est151:\*  
221: gb\_est152:\*  
222: gb\_est153:\*  
223: gb\_est154:\*  
224: gb\_est155:\*  
225: gb\_est156:\*  
226: gb\_est157:\*  
227: gb\_est158:\*  
228: gb\_est159:\*  
229: gb\_est160:\*  
230: gb\_est161:\*  
231: gb\_est162:\*  
232: gb\_est163:\*  
233: gb\_est164:\*  
234: gb\_est165:\*  
235: gb\_est166:\*  
236: gb\_est167:\*  
237: gb\_est168:\*  
238: gb\_est169:\*  
239: gb\_est170:\*  
240: gb\_est171:\*  
241: gb\_est172:\*  
242: gb\_est173:\*  
243: gb\_est174:\*  
244: gb\_est175:\*  
245: gb\_est176:\*  
246: gb\_est177:\*  
247: gb\_est178:\*  
248: gb\_est179:\*  
249: gb\_est180:\*  
250: gb\_est181:\*  
251: gb\_est182:\*  
252: gb\_est183:\*  
253: gb\_est184:\*  
254: gb\_est185:\*  
255: gb\_est186:\*  
256: gb\_est187:\*  
257: gb\_est188:\*  
258: gb\_est189:\*

pred No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB  | ID        | Description |
|------------|-------|-------------|--------|-----|-----------|-------------|
| 1          | 467.4 | 9.8         | 573    | 139 | BE755695  | 209602 MA   |
| 2          | 449.2 | 9.4         | 573    | 108 | AU135340  | AU135340    |
| 3          | 398.8 | 8.4         | 482    | 138 | BE650938  | BE650938    |
| 4          | 371.2 | 7.8         | 954    | 134 | BE479794  | BE479794    |
| 5          | 350.8 | 7.4         | 515    | 110 | AV748467  | AV748467    |
| 6          | 305.6 | 6.4         | 368    | 136 | BE478981  | BE478981    |
| 7          | 263.8 | 5.5         | 517    | 17  | AI176907  | AI176907    |
| 8          | 250.2 | 5.3         | 466    | 24  | AI1760509 | AI1760509   |
| 9          | 249   | 5.2         | 326    | 134 | BE458694  | BE458694    |
| 10         | 240   | 5.0         | 308    | 134 | BE449764  | BE449764    |
| 11         | 239.8 | 5.0         | 505    | 19  | AI140862  | AI140862    |
| 12         | 236.2 | 5.0         | 484    | 148 | BE420450  | BE420450    |
| 13         | 217.4 | 4.6         | 401    | 141 | BE847128  | BE847128    |
| 14         | 216.2 | 4.5         | 314    | 128 | BE218602  | BE218602    |
| 15         | 201.4 | 4.2         | 422    | 17  | AI233991  | AI233991    |
| 16         | 201.4 | 4.2         | 444    | 116 | AW489638  | AW489638    |
| 17         | 200   | 4.2         | 400    | 22  | AI602045  | AI602045    |
| 18         | 194   | 4.1         | 265    | 3   | AA184901  | AA184901    |
| 19         | 193.4 | 4.1         | 384    | 141 | BE852200  | BE852200    |
| 20         | 193.2 | 4.1         | 247    | 116 | AW446518  | AW446518    |
| 21         | 192.4 | 4.0         | 313    | 128 | BE219780  | BE219780    |
| 22         | 190.2 | 4.0         | 590    | 166 | BE38063   | BE38063     |
| 23         | 190   | 4.0         | 431    | 116 | AW456831  | AW456831    |
| 24         | 174   | 3.7         | 334    | 163 | BE097510  | BE097510    |
| 25         | 172.2 | 3.6         | 383    | 150 | BE522383  | BE522383    |
| 26         | 158.8 | 3.3         | 927    | 106 | AL552613  | AL552613    |
| 27         | 158.4 | 3.3         | 3739   | 192 | AK014835  | AK014835    |
| 28         | 157   | 3.3         | 397    | 226 | AQ270160  | AQ270160    |
| 29         | 156.2 | 3.3         | 683    | 243 | A426543   | A426543     |
| 30         | 152   | 3.2         | 497    | 15  | AI072568  | AI072568    |
| 31         | 150.2 | 3.2         | 286    | 135 | BB475629  | BB475629    |
| 32         | 136.6 | 2.9         | 745    | 106 | AL577036  | AL577036    |
| 33         | 136.2 | 2.9         | 965    | 146 | BE302800  | BE302800    |
| 34         | 133.2 | 2.8         | 918    | 106 | AL556703  | AL556703    |
| 35         | 131.6 | 2.8         | 864    | 18  | AI132453  | AI132453    |
| 36         | 131.6 | 2.8         | 807    | 108 | AI134073  | AI134073    |
| 37         | 131.2 | 2.8         | 741    | 155 | BE568567  | BE568567    |
| 38         | 130.6 | 2.7         | 942    | 174 | BE175139  | BE175139    |
| 39         | 130   | 2.7         | 657    | 17  | AI173322  | AI173322    |
| 40         | 130   | 2.7         | 759    | 154 | BE533260  | BE533260    |
| 41         | 129   | 2.7         | 1038   | 141 | BE911604  | BE911604    |
| 42         | 127.8 | 2.7         | 290    | 141 | BE848985  | BE848985    |
| 43         | 127.2 | 2.7         | 775    | 175 | BE246010  | BE246010    |
| 44         | 127   | 2.7         | 857    | 175 | BE242396  | BE242396    |
| 45         | 125.4 | 2.6         | 608    | 137 | BE554133  | BE554133    |

## ALIGNMENTS

| RESULT     | 1  | ALIGNMENTS  |
|------------|--|-------------|
| LOCUS      | BE755695   | 573 bp mRNA |
| DEFINITION | 209602 MARC 2B0V Bos taurus CDNA 5', mRNA sequence.  | EST         |
| ACCESSION  | BE755695   | 15-SEP-2000 |
| VERSION    | BE755695.1   | GI:10169687 |
| KEYWORDS   | EST.   |             |
| SOURCE     | COW.   |             |
| ORGANISM   | Bos taurus   |             |
| REFERENCE  | Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Groesse,W.M., Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Laegreid ,W.W. and Keele,J.W. |             |
| AUTHORS    | 1 (Bases 1 to 573)   |             |
| TITLE      | Design and use of four pooled tissue normalized cDNA libraries for   |             |

## JOURNAL COMMENT

EST discovery in cattle

Unpublished (2000)

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt-trimmed with phred

v0.980904.e. Vector identified by cross-match with the -minscore 18

and -mismatch 12 options.

PCR primers

FORWARD: AGCAACAGCTATGACCAT

BACKWARD: GTTTCGCCAGTCACGACG

Plate: 59 row: N column: 5

Seq primer: ATTAGGTGACCTATAG.

Location/Qualifiers

1. 573

/organism="Bos taurus"

/db\_xref="taxon:9913"

/clone\_lib="MARC 2B0V"

/lisse\_type="pooled"

/lab\_host="DH10B"

/note="Vector: pCMV SPORTe; Site\_1: XbaI; Site\_2: XhoI;

Library made from pooled tissue from testis, thymus,

semiteadonous muscle, longissimus muscle, pancreas,

adrenal, and endometrium."

BASE COUNT 149 a 134 c 146 g 144 t

## ORIGIN

Query Match 9 88; Score 467.4; DB 139; Length 573;

Best Local Similarity 88.5%; Pred. No. 2.3e-124;

Matches 507; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

|    |      |  |      |
|----|------|--|------|
| Qy | 3358 | ttcaactgaataatgagaagaatcgaagcgtcccgcaatatcagatggaagatccc     | 3417 |
| Db | 1    | TTTGGCCGAAACATGCGAAGAACTGGGGCGCCCGCATGCGACGACGAGACCT         | 60   |
| Qy | 3418 | actttaagaagaattatcgcttcacatgcaataatgagctacataatgatacactct    | 3477 |
| Db | 61   | ACTTTTCAAGAAAGATATCCCTTCATCATCATATGCTACGTGATGATACCTCCT       | 120  |
| Qy | 3478 | ggcttaagatgctgaagatcgaagatcgaatgcatcgtcgaacatggagacgaat      | 3537 |
| Db | 121  | GGCTTATGATGCTCGACATCAAGGATTAAGGATGCTGCTCAGCATGCGACGAAT       | 180  |
| Qy | 3538 | gaaacatcattcattcattcattcattcattcattcattcattcattcattcattcatt  | 3597 |
| Db | 181  | GAAATATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT      | 240  |
| Qy | 3598 | tataaattgacatgtaacatctcattcattcattcattcattcattcattcattcatt   | 3657 |
| Db | 241  | TATTAATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT        | 300  |
| Qy | 3658 | tccaaagctggaatttggcggtggaatggccttatttggcggaatctacatgctggatg  | 3717 |
| Db | 301  | TCCAAAGGTTGGGACTTGGCGGATTAAGTGTATTGGCGACCTCAAGCTGGGATG       | 360  |
| Qy | 3718 | agcaactttcttggtgtacgaataaagtgtaagatccctctggaagatgcttcgga     | 3777 |
| Db | 361  | AGCACTCTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT       | 420  |
| Qy | 3778 | caactagagatttcaagattcaagcttcaggaacataatgagagtggtgcccagaagctg | 3837 |
| Db | 421  | CGCATTAAGATTTTCAGATTACAGCTTCAGCAATATGAGACAGTGGCCCCCAAGCTG    | 480  |
| Qy | 3838 | ggcagacttatattcccggaataatgagcttgagagcgaagagcccttctcttg       | 3897 |
| Db | 481  | GCCAGACTTATTTATTTCTGGATCAATCAACGCTGGGACCAAGGATTCCTCTTGG      | 540  |
| Qy | 3898 | atcaagtgatcgtgttgacacatgatatc 3930                           |      |



| DB         | 541  | ATCAGGTCGATCTGTTGGCGCCGATGATTAT | 573      |
|------------|--|---------------------------------|----------|
| RESULT     | 2  |                                 |          |
| LOCUS      | AU135340   |                                 |          |
| DEFINITION | AU135340 PLACE1 Homo sapiens CDNA clone PLACE1001786 5', mRNA sequence.  |                                 |          |
| ACCESSION  | AU135340   |                                 |          |
| VERSION    | AU135340.1   | GI:10995879                     |          |
| KEYWORDS   | EST.   |                                 |          |
| SOURCE     | human.   |                                 |          |
| ORGANISM   | Homo sapiens   |                                 |          |
| REFERENCE  | Okuyama, M.; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 815)  |                                 |          |
| AUTHORS    | Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, Y., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and Isogai, T.   |                                 |          |
| TITLE      | HRI human CDNA project   |                                 |          |
| JOURNAL    | Unpublished (2000)   |                                 |          |
| COMMENT    | Contact: Takao Isogai<br>Genomics Laboratory<br>Helix Research Institute<br>1532-3 Yana, Kisarazu, Chiba 292-0812, Japan<br>Tel: 81-438-52-3851<br>Fax: 81-438-52-3952<br>Email: genomics@hri.co.jp<br>HRI human CDNA project; 5'- & 3'-end one pass sequencing; Helix Research Institute; CDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute. |                                 |          |
| FEATURES   | Location/Qualifiers  |                                 |          |
| Source     | 1..815<br>/organism="Homo sapiens"<br>/db_xref="taxon:9606"<br>/clone="PLACE1001786"<br>/clone_lib="PLACE1"<br>/tissue_type="placenta"<br>/note="Vector: pME18SFL3"  |                                 |          |
| BASE COUNT | 200 a 207 c 204 g 197 t  |                                 |          |
| ORIGIN     |  |                                 | 7 others |

| Query Match           | 9.4%   | Score 449.2        | DB 108       | Length 815 |
|-----------------------|--|--------------------|--------------|------------|
| Best Local Similarity | 98.3%  | Pred. No. 5.4e-119 |              |            |
| Matches 454           | Conservative   | 0                  | Mismatches 8 | Indels 0   |
|                       |  |                    |              | Gaps 0     |
| QY 4057               | atgagctctcttggcgaatgtagattcatcttggaataaacaacaaattttaacctcca      | 4116               |              |            |
| Db 92                 | AAGGCTCTTCTTTGGCAATGTGGATTCTATCTGGATTAACACAAATATTTTAACTCCCA      | 151                |              |            |
| QY 4117               | attatctgcgaatacatccgtcttgcaaccaactcatatagaatccgacactctgcg        | 4176               |              |            |
| Db 152                | ATTATTGCTCATATACATCCGTTTGCAACCAACTCATTTAGCATTTGCGACACTCTTGGC     | 211                |              |            |
| QY 4177               | atggaatctgatatggcgctgtgatttaataatagtctgcaatctccaatggaaatggaataaa | 4236               |              |            |
| Db 212                | ATGGAGTTGATGGCGCTGATGATTTAAATATTGCAGATGCCAFTTGGGATGGGAGATTA      | 271                |              |            |
| QY 4237               | gcaataatcagaatgacaagaattactgcttcatctacttaccataatgtttgccacttg     | 4296               |              |            |
| Db 272                | GCAATATTCAGATGCACAGATTACTGTTTACTCTTACTTTACCAATATGTTTGCCACTGG     | 331                |              |            |
| QY 4297               | tctctctcaaaaagctcgacttcaacctccaagaagagagtaatgactcggaacctcaagt    | 4356               |              |            |
| Db 332                | TCTCTTCAAAAAGCTCGAAGTTTCACTGCCCAAGGAGAGTAATGCTCGAGACTCTCAGTGG    | 391                |              |            |
| QY 4357               | aataatccaaaagatlggcttgcaatggaatctccagaagaacaaatgaagaatccacagagta | 4416               |              |            |
| Db 392                | AATATTCAAAAGAGTGGCTGCAAGTGACCTTCAGAAACACATGAATCAACAGAGATTA       | 451                |              |            |
| QY 4417               | actactcaggaatgaataatctctgcttaccagaatgtatgtgaagaagttctctcatctcc   | 4476               |              |            |

|            |   |  |             |
|------------|---|--|-------------|
| Db         | 452   | ACTGCTAGGAGAGTAAATCTCTTCGTACCAAGATATATGTANAGAGATTCTCATCTCC | 511         |
| OY         | 4477  | agcagctcaagatgycacatcgatggactcttttccagaat                  | 4518        |
| Db         | 512   | AGCAGTCAAGATGGCCATCATGATGCCTCGCTACTCTCGAATT                | 553         |
| RESULT     | 3   |  |             |
| LOCUS      | BE650938  |  |             |
| DEFINITION | BE650938  | 482 bp mRNA  | EST         |
| ACCESSION  | U1-M-BH3-asm-g-04-0-0-I .I1 NIH_BMAP_M_S4   | Mus musculus cDNA clone                                    | 06-SEP-2000 |
| VERSION    | U1-M-BH3-asm-g-04-0-0-I 5'  | mRNA sequence.   |             |
| KEYWORDS   | BE650938  |  |             |
| SOURCE     | BE650938.1  | GI:9976762   |             |
| ORGANISM   | EST.  |  |             |
|            | house mouse.  |  |             |
|            | Mus musculus  |  |             |
|            | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;   |  |             |
|            | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  |  |             |
| REFERENCE  | 1 (bases 1 to 482)  |  |             |
| AUTHORS    | Bonaldo,M.F., Lennon,G. and Soares,M.B.   |  |             |
| TITLE      | Normalization and subtraction: two approaches to facilitate gene discovery  |  |             |
| JOURNAL    | Genome Res.   | 6 (9), 791-806   | (1996)      |
| MEDLINE    | 9704477   |  |             |
| COMMENT    | Contact: Chin, H<br>National Institute of Mental Health<br>6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD<br>20892-9643, USA<br>Tel.: 301 443 1706<br>Fax: 301 443 9890<br>Email: mestr@nhi.nih.gov<br>CDNA Library Preparation: M.B. Soares Lab Clone distribution:<br>Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It<br>should be noted that Bento Soares is generating a small number of<br>additional specialized non-redundant arrays of BMAP cDNAs whose<br>availability will be considered under appropriate and limited<br>collaborative arrangements<br>Seq primer: M13 Reverse.<br>Location/Qualifiers<br>1..482 |  |             |
| FEATURES   |   |  |             |
| source     |   |  |             |

```

Email: mst@email.nih.gov
CDNA Library Preparation: M.B. Soares Lab Clone distribution:
Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
should be noted that Bento Soares is generating a small number of
additional specialized non-redundant arrays of BMAP cDNAs whose
availability will be considered under appropriate and limited
collaborative arrangements
Seq primer: M13 Reverse.

Location/Qualifiers
1. 482
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BH3-asm-g-04-0-UI"
/clone_lib="NIH_BMAP_M.S4"
/dec_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/vector="pR173D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The
NIH_BMAP_M.S4 library is a subtracted library of a series,
ultimately derived from a mixture of individually tagged
normalized libraries from ten regions of the mouse brain
(cerebellum, brain stems, olfactory bulbs, hypothalamus,
cortex, amygdala, basal ganglia, pineal gland, striatum,
hippocampus) after a series of subtractions to reduce the
representation of cDNAs from which ESTs had already been
generated. The following serially subtracted libraries
were generated in this process: NIH_BMAP_M.S4,
NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, NIH_BMAP_M.S3.1,
NIH_BMAP_M.S2, NIH_BMAP_M.S1. The subtracted library
(NIH_BMAP_M.S4) was constructed as follows: PCR amplified
cDNA inserts from NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, and
NIH_BMAP_M.S3.1 clones from which 3' ESTs had been derived
was used as a driver in a hybridization with a pool of
the NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, and NIH_BMAP_M.S3.1
libraries in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library)
was purified by hydroxyapatite column chromatography
converted to double-stranded circles and electroporated
into DH10B bacteria (Life Technologies) to generate the
NIH_BMAP_M.S4 library. This procedure has been previously

```



|                           |  |
|---------------------------|--|
| RESULT                    | 5  |
| LOCUS                     | AV748467 515 bp mRNA EST 19-OCT-2000   |
| DEFINITION                | AV748467 NPC Homo sapiens cDNA clone NPCXAA05 5', mRNA sequence.   |
| ACCESSION                 | AV748467   |
| VERSION                   | AV748467.1 GI:10906315   |
| KEYWORDS                  | EST.   |
| SOURCE                    | human.   |
| ORGANISM                  | Homo sapiens<br>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;<br>Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.<br>1 (bases 1 to 515)   |
| REFERENCE                 | Song,H., Peng,Y., Gu,Y., Yang,Y., Gao,G., Xiao,H., Xu,X., Li,N.,<br>Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., G,<br>'W., Ju,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Ye,M., Zhang,Q.<br>, Han,Z., Chen,Z., Hu,R. and Chen,J.<br>Homo sapiens NPC library cDNA clones<br>Unpublished (2000)<br>Contact: Qionghua Zhang<br>Shanghai Institute of Endocrinology, Rui-Jin Hospital<br>197 Rui-Jin II Road, Shanghai 200025, P. R. China<br>Tel.: 86-21-64370045(ex.663332)<br>Fax: 86-21-64743206<br>Email: mbsheims.stn.sh.cn<br>This clone is available at Shanghai Hematology Institute in<br>Shanghai.<br>Chinese National Human Genome Center at Shanghai<br>351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong.<br>Location/Qualifiers<br>1..515<br>/organism="Homo sapiens"<br>/db_xref="taxon:9606"<br>/clone="NPCXA05"<br>/clone_1lb="NPC"<br>/tissue_type="pituitary"<br>/dev_stage="Adult"<br>/lab_host="SOLR"<br>/note="Vector: pBluescript sk(-); Site.1: EcoRI; Site.2:<br>XhoI" |
| FEATURES                  | source   |
| BASE COUNT                | 132 a 114 c 131 g 135 t 3 others   |
| ORIGIN                    |  |
| Query Match               | 7.4%; Score 350.8; DB 110; Length 515;   |
| Best Local Similarity     | 95.3%; Pred. No. Ze-90;  |
| Matches 383; Conservative | 0; Mismatches 14; Indels 5; Gaps 2;  |
| OY 4057                   | atgagcttccttggaagtgtgatcatcattgtggagaataaacacaatatcttaacctcca 4116   |
| Db 111                    | AAGGCTCTTGTGGCAATGTGGATTTCATCTGGATTAACAACAATAATTTTAACTCCA 170  |
| OY 4117                   | attatgtcgatatacatccglttgcaaccacatcatlatagcatlccgcagcaactctgcg 4176   |
| Db 171                    | ATTATGTGTCGATTACATCGCTTGCCAGCACCACCATCATTAATAGCATTCGCAGCACTCTGCC 230   |
| OY 4177                   | atcgagctgtgatggcgctgtgattaaatagttgcagcatgcattgaggaaatgagagtaa 4236   |
| Db 231                    | ATGGATGTGAATGGCGCTTGCAATTAATAGTTACACATCCCATTTGGGAATGAGAGTAAA 290   |
| OY 4237                   | gcaataatagatggcacagattactctctcatctctacttacaatatgtttgccaccttg 4296  |
| Db 291                    | GCAATATCAGATGACACAGATTACTCTTCATCTCACTTTAACCAATAATGTTCACCACTTG 350  |
| OY 4297                   | tctccttcaaaaagctcgacttcaacttcaactccaaggagagatgaatgccttggagacctcagtg 4356   |
| Db 351                    | TCTCTTCAAAAAGCTCGAATTCACCTCTCAAGSAGSAGTAATGCTCGAGACTCAAGTG 410   |
| OY 4357                   | aataatccaaaaagatlgcgtcgaatlgagatccacagaagaacaatgaagaatlcacagagta 4416  |
| Db 411                    | AATATCCAAAAAGATGTGGCTTCAAGTGACCTTCAGAAACACATATNGTGCACA-NGAAGTA 469   |
| OY 4417                   | actactcaggagatataaatctctgcttaccagcatgtagtg 4458  |

| DB                        | 470   | ACTACTTAAGAGTA---CTCTTCTTACCAGCATGTATGTCG                       | 507         |
|---------------------------|---|---|-------------|
| RESULT                    | 6   |   |             |
| LOCUS                     | BE478981  | 368 bp  | mRNA        |
| DEFINITION                | 163493 BARC 5BOV Bos taurus cDNA 5', mRNA sequence.   |   | EST         |
| ACCESSION                 | BE478981  |   | 28-AUG-2000 |
| VERSION                   | BE478981.1  | GI:9598514  |             |
| KEYWORDS                  | EST.  |   |             |
| SOURCE                    | COW.  |   |             |
| ORGANISM                  | Bos taurus  |   |             |
| REFERENCE                 | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.  |   |             |
| AUTHORS                   | 1 (bases 1 to 368)<br>Sonstegard,T.S., Capuco,A.V., Van Tassel,C.P., Ashwell,M.S. and Wells,K.D.  |   |             |
| TITLE                     | Mapping of Expressed Sequence Tags from a normalized bovine mammary gland cDNA library  |   |             |
| JOURNAL                   | Unpublished (2000)  |   |             |
| COMMENT                   | Contact: Sonstegard TS<br>USDA, ARS, Beltsville Agricultural Research Center<br>Bldg. 200 Rm 2A, Beltsville, MD 20705, USA<br>Tel: 301 504 8416<br>Fax: 301 504 8414<br>Email: tads@sls.barc.usda.gov<br>Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -m1nscore 18 and -minmatch 12 options.<br>PCR primers<br>FORWARD: AGGAACAGCTATGACCAT<br>BACKWARD: GTTTCCTCCAGTCACGAG<br>Plate: 141 row: G column: 5<br>Seq primer: ATTAGTGCACATATAC.<br>Location/Qualifiers<br>1..368<br>/organism="Bos taurus"<br>/db_xref="taxon:9913"<br>/clone_lib="BARC 5BOV"<br>/tissue_type="pooled"<br>/lab_host="DH10B"<br>/note="Vector: pCMV SPORT6, Site_1: XbaI, Site_2: XhoI; Library made from pooled mRNA isolated from mammary tissues at eight physiological, developmental, and disease states." |   |             |
| FEATURES                  | source  |   |             |
| BASE COUNT                | 90 a 87 c 98 g 93 t   |   |             |
| ORIGIN                    |   |   |             |
| Query Match               | 6.4%; Score 305.6; DB 136; Length 368;  |   |             |
| Best Local Similarity     | 89.4%; Pred. No.2.5e-77;  |   |             |
| Matches 329; Conservative | 0; Mismatches 39; Indels 0; Gaps 0;   |   |             |
| QY                        | 3598  | tataaaatgacacgttacaatctctatccaggctgtttttagagacagtggaaatgtacc    | 3657        |
| Db                        | 1   | TGTAAATGGCGCGTCTACAACTCTACCCAGGTGCTTTGAGACGCTGGAATGCTACCA       | 60          |
| QY                        | 3658  | tccaaagctggaatttggcgggtggaatgaagcctatttggcggagacactcaatgctggagt | 3717        |
| Db                        | 61  | TCCAAGGTTGGGACTTGGCGGATTAAGTCTTATTTGGGAGACACTCAACCTGGGATG       | 120         |
| QY                        | 3718  | agcaacactttctctgtgtgtacagcaataatgtctagactccctctggatggtctctga    | 3777        |
| Db                        | 121   | AGCACTCTCTTCTGTGTGTGTACAGCAAGAGTGTCAAAATTCACCTGGGAATGGCTTCTGA   | 180         |
| QY                        | 3778  | caacatagagatttccagattacagcttcctcaagacataatgacagtggtggcccaagctg  | 3837        |
| Db                        | 181   | CGCATTTAGAGATTTTCAATATTACACTTCAGAGACAATATGACACTGGGCCCCCAAGCTG   | 240         |
| QY                        | 3838  | gccaagctcatattatccggatcaataaagtcctgtgaagacccaaggaacccctttcttgg  | 3897        |

Db 241 GCCAGCTCATTTATCTGTGATCATCAACGCCGTGGAGCACCACCAAGATCCCTCTCTGG 300

Qy 3898 atcaaggtgagctctgttggcaccatgatattatcaagcatcaacaccagtgccgt 3957  
|||||  
Db 301 ATCAAGGTGATCTGTGGCCCATGATTTATTCACAGACATCTGACTCAAGGTGCCCG 360

Qy 3958 cagaagtt 3965  
|||||  
Db 361 CAGAAGTT 368

RESULT 7  
A1176907/c 517 bp mRNA EST 20-JAN-1999  
LOCUS EST220513 Normalized rat ovary, Bento Soares Rattus sp. cDNA clone  
DEFINITION ROVER42.3' end, mRNA sequence.  
ACCESSION A1176907  
VERSION A1176907.1 GI:3727545  
KEYWORDS EST.  
SOURCE Rattus sp.  
ORGANISM Rattus sp.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 517)  
Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,  
Kerlavage,A.R. and Adams,M.D.  
Rat Genome Project: Generation of a Rat EST (RESt) Catalog & Rat  
Gene Index  
Unpublished (1998)  
JOURNAL Contact: Lee, NH  
COMMENT The Institute for Genomic Research  
9712, Medical Center Drive, Rockville, MD 20850, USA  
Tel: (301)-838-3529  
Fax: (301)-838-0208  
Email: nhlee@ligr.org  
Seq primer: M13-21.  
FEATURES  
source location/Qualifiers  
1..517  
/organism="Rattus sp."  
/db\_xref="ATCC (Inhost):2031594"  
/db\_xref="taxon:10118"  
/clone="ROVER42"  
/clone\_lib="Normalized rat ovary, Bento Soares"  
/note="Organo: ovary; Vector: pT7T3Pac; Site\_1: EcoRI;  
Site\_2: NotI"

BASE COUNT 93 a 174 c 154 g 96 t

ORIGIN

Query Match 5.5%; Score 263.8; DB 17; Length 517;  
Best Local Similarity 70.9%; Pred. No. 4.2e-65;  
Matches 365; Conservative 0; Mismatches 147; Indels 3; Gaps 1;

Qy 2743 cagagccccgagcttcaagaagaacaacagacatttatttctgtcagtgagag 2802  
|||||  
Db 513 CAGGAGCCCGCGGCTTTCAGAGAAACGCGACTATTTCATCGCGCGGAAGAG 454

Qy 2803 ctctggatattggatagatagctcccaagtttcaagaagagctcagatggc 2862  
|||||  
Db 453 CTTTGGACATACGGGGGCGCTGTGCTCGAGCTCCGGGAGACAGGGCGGTGAGTGG 394

Qy 2863 agtgcctcaagttcaagaagtttccagaatttactgtatgctccttaccag 2922  
|||||  
Db 393 GAGCGACTAGGTTCCGGAAGTGTGTTTCCGTGAGTTTCACGAGCGGTGTTCAACG 334

Qy 2923 cccctataacgctggagaactaatggaacttgggaactccctggggccatataagaagca 2982  
|||||  
Db 333 CGCGTGACCGGTGGAGCTGGAAGACACCTGGGGCTACTGGGGCCATACATACGGGG 274

Qy 2983 gaagtgaagataataatcatgtaacttcaagaatcagcctctgctccttccctc 3042  
|||||  
Db 273 GAAGTGGAGACAAACATCGTGTGACCTTCGGAACACAGGCTCCGCCCTACTCCTTC 214

Qy 3043 tatctagccctatttcttaaggaagatcagaagcgaagagacagactagaanaaac 3102  
|||||  
Db 213 TACTCCAGCCTCGTGTCTTACCCGGAAGCCGACGCT---GGAGCTGACCAAGAGCAAC 157

Qy 3103 ttgtcaagcctaalgaaacaaacttacttcttgaagtgacaacatcatatggcacc 3162  
|||||  
Db 156 TTGTGGCGGCCAACAGACCAAAACTTACTTCTGGAGAGTCCGGCCCAACATGGGCCCC 97

Qy 3163 actaagaatgatttgactgcaagcctgggttatttctctgtatgttgcctggaanaa 3222  
|||||  
Db 96 ACGGACGCGCAGTTTGACTGCAAGGCCCTGATTTCTCCGACGTGACCTGACGCA 37

Qy 3223 gatgtgacccagcagctgattgacccctctggt 3257  
|||||  
Db 36 GACCTGCACCTGGGGCTGATTGGGCCACTGCTGCT 2

RESULT 8  
A1760509/c 466 bp mRNA EST 20-DEC-1999  
LOCUS wh64d04.x1 NCLCGAP\_Kid11 Homo sapiens cDNA clone IMAGE:2385511 3'  
DEFINITION similar to gb:M14113 COAGULATION FACTOR VIII PRECURSOR (HUMAN);,  
mRNA sequence.  
ACCESSION A1760509  
VERSION A1760509.1 GI:5176176  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 466)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/cgi/cgap  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgaps-remail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
DNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www.bio.lnl.gov/bdrp/image/image.html  
Insert Length: 1564 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 404.  
FEATURES  
source location/Qualifiers  
1..466  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2385511"  
/clone\_lib="NCI-CGAP\_Kid11"  
/lab\_host="DH10B"  
/note="Organo: kidney; Vector: pT7T3P-Pac (Pharmacia) with  
a modified polylinker. Site\_1: Not I; Site\_2: Eco RI;  
Plasmid DNA from the normalized library NCI-CGAP\_Kid3 was  
prepared, and ss circles were made in vitro. Following HAP  
purification, this DNA was used as tracer in a subtractive  
hybridization reaction. The driver was PCR-amplified cDNAs  
from a pool of 5,000 clones made from the same library  
(cloneIDs 1322376-1323911, 1456007-1456775, and  
1500352-1502855). Subtraction by Bento Soares and M.  
Fatima Bonaldo."

BASE COUNT 128 a 96 c 77 g 165 t

ORIGIN

Query Match 5.3%; Score 250.2; DB 24; Length 466;  
Best Local Similarity 95.2%; Pred. No. 3.7e-61;  
Matches 258; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

| DB         | Accession  | Version    | Keywords | Source | Organism    | Reference | Authors |
|------------|------------|------------|----------|--------|-------------|-----------|---------|
| OY         | 1466       |            |          |        |             |           |         |
| DB         | 453        |            |          |        |             |           |         |
| OY         | 1526       |            |          |        |             |           |         |
| DB         | 393        |            |          |        |             |           |         |
| OY         | 1586       |            |          |        |             |           |         |
| DB         | 333        |            |          |        |             |           |         |
| OY         | 1646       |            |          |        |             |           |         |
| DB         | 273        |            |          |        |             |           |         |
| OY         | 1706       |            |          |        |             |           |         |
| DB         | 213        |            |          |        |             |           |         |
| RESULT     | 9          |            |          |        |             |           |         |
| LOCUS      | BB458694   |            |          |        |             |           |         |
| DEFINITION | BB458694   | 326 bp     | mRNA     | EST    | 21-JUN-2000 |           |         |
| ACCESSION  | BB458694   |            |          |        |             |           |         |
| VERSION    | BB458694.1 | GI:9354187 |          |        |             |           |         |
| KEYWORDS   |            |            |          |        |             |           |         |
| SOURCE     |            |            |          |        |             |           |         |
| ORGANISM   |            |            |          |        |             |           |         |
| REFERENCE  |            |            |          |        |             |           |         |
| AUTHORS    |            |            |          |        |             |           |         |
| TITLE      |            |            |          |        |             |           |         |
| JOURNAL    |            |            |          |        |             |           |         |
| COMMENT    |            |            |          |        |             |           |         |

| FEATURES   |   | Source  |                     |
|--|---|---|---------------------|
| Location/Qualifiers  |   | 1. .326   |                     |
| /organism="Mus musculus"   |   | /db_xref="taxon:10090"  |                     |
| /clone="D130058E20"  |   | /clone_lib="RIKEN full-length enriched, 12 days embryo  |                     |
| /tissue_type="spinal ganglion"                                     |   | /dev_stage="12 days embryo"   |                     |
| /lab_host="DH10B"  |   | /note="Site.1: SalI; Site.2: BamHI. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer 15' |                     |
| GAGGAGAGAGCGGCGCGAATCGAGTTTCTTTTCTTTTNN 3']                        |   | cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence 15'   |                     |
| GAGGAGAGATTTCGAGTTATTTATTTATTTATCCCGCCCCCCC 3']                    |   | cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLX I."  |                     |
| BASE COUNT   |   | 87 a 82 c 56 g 101 t  |                     |
| ORIGIN   |   |   |                     |
| Query Match  | 5.2%  | Score 249;  | DB 134; Length 326; |
| Best Local Similarity  | 86.0%;  | Fred. No. 7.1e-61;  |                     |
| Matches 276;   | Conservative 0;   | Mismatches 45;  | Indels 0; Gaps 0;   |
| Qy 1009  | tggtctaaatgcacacagtcagtgtatgtgaacagtgctctgcgaagtcgtattgga       | 1068  |                     |
| Db 2   | tgcctctttttatccacacttttcgaagccactttacacagctcttttccagctctgattgga | 61  |                     |
| Qy 1069  | tgccacaggaatcagctctatgtgcatgtgatgtgaatggagcaccactctgaatgac      | 1128  |                     |
| Db 62  | taccataggaaatcagctctgactgacgacgagttggaatggcgtccacgctgaattacac   | 121   |                     |
| Qy 1129  | tcaatattccctcgaaggtcacaacattctcttggaaggaacacatcgccagcgctcttgaa  | 1188  |                     |
| Db 122   | tcaaatattccctcgaaggtcacaacatttttggaggaacacacccctcaagcttcattggag | 181   |                     |
| Qy 1189  | atctgcacaataactctctctactctgctcaaacacacttgatgaccttggaacagttcta   | 1248  |                     |
| Db 182   | atattacacacattactcttcttctgctcaaacacactcttgatgaccttggaacagttcta  | 241   |                     |
| Qy 1249  | ctgttttgcatalatctctccacacaacatgattgcatgtaagcttattcctcaagtagac   | 1308  |                     |
| Db 242   | ctattttttttcattttctttcccataaacatgatgtgcatgattggtggcagttccta     | 301   |                     |
| Qy 1309  | agctgtccaagaggaaccccaa  | 1329  |                     |
| Db 302   | agctgcctctgaggaattcccaa   | 322   |                     |
| RESULT 10  |   |   |                     |
| LOCUS  |   |   |                     |
| BB449764 308 bp mRNA EST 21-JUL-2000                               |   |   |                     |
| DEFINITION   |   |   |                     |
| BB449764 RIKEN full-length enriched, 9 days embryo Mus musculus    |   |   |                     |
| cDNA clone DO30073H24.3 similar to I05573 Mus domesticus           |   |   |                     |
| consolation factor VII mRNA, mRNA sequence.                        |   |   |                     |
| ACCESSION  |   |   |                     |
| BB449764   |   |   |                     |
| VERSION  |   |   |                     |
| BB449764.1 GI:9310799  |   |   |                     |
| KEYWORDS   |   |   |                     |
| EST.   |   |   |                     |
| house mouse.   |   |   |                     |
| SOURCE   |   |   |                     |
| Mus musculus   |   |   |                     |
| ORGANISM   |   |   |                     |
| Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  |   |   |                     |
| Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. |   |   |                     |



QY 1185 ggaatctgcgaataacttcttactgtctcaaacactcttgatgagccttgagcaatc 1244  
          |||||          |||||          |||||          |||||          |||||  
Db 325 GGAGATCTCACATCTATGACTGTCTACCGCTCAGACGTGGATGGAGCTGGGGCGGT 266  
          |||||          |||||          |||||          |||||          |||||  
QY 1245 tctactgtttgcatctcttccaccacatgatgagcttgtaagcttggtaaatc 1304  
          |||||          |||||          |||||          |||||          |||||  
Db 265 CCTGCTCTTGTGTACATCTCTCCACACAGACAGGTGGTGAAGACCTTACGTCCAGT 206  
          |||||          |||||          |||||          |||||          |||||  
QY 1305 agacagctgtccagaggaaccccaactcgaatgaaaaataatgaagaagcgaagaacta 1364  
          |||||          |||||          |||||          |||||          |||||  
Db 205 GGACAACTGCCCGCAGCGCCCGCCAGGGGAGGGAAGGAGCAGACAGACAGACTA 146  
          |||||          |||||          |||||          |||||          |||||  
QY 1365 tgaatgatctctactcactcgaatgagatgagtgagtgaggttgatgatacactctcc 1424  
          |||||          |||||          |||||          |||||          |||||  
Db 145 TGATGA-----TCTGGACTCTGGAATGAGACTTTCACCTGGGACCCCGATGCCGCC- 91  
          |||||          |||||          |||||          |||||          |||||  
QY 1425 ttccttatccaattgtctgaagtgtccaagaagccttaacttggatcattacat 1484  
          |||||          |||||          |||||          |||||          |||||  
Db 92 -CCTTGTCCAGGTCCGCTCAGACGCCAAGACACCCCAAGACTGTGGTGCATCACT 35  
          |||||          |||||          |||||          |||||          |||||  
QY 1485 tgcctcgaagagagagagactgagactatgctcc 1517  
          |||||          |||||          |||||          |||||          |||||  
Db 34 CGCAGCAGAGAGGGGGGAGCTGGACTACGCCGC 2  
          |||||          |||||          |||||          |||||          |||||  
RESULT 12  
BF420450 484 bp mRNA EST 28-NOV-2000  
LOCUS UI-R-BJ2-bpw-f-03-0-UI.s1 UI-R-BJ2 Rattus norvegicus cDNA clone  
DEFINITION UI-R-BJ2-bpw-f-03-0-UI 3', mRNA sequence.  
ACCESSION BF420450 GI:11408479  
VERSION BF420450.1  
KEYWORDS EST.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus.  
          Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
          Rattus.  
REFERENCE 1 (bases 1 to 484)  
AUTHORS Bernaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene  
          discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 9704477  
COMMENT Contact: Soares, MB  
          Program for Rat Gene Discovery and Mapping  
          University of Iowa  
          451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
          Tel: 319 335 8250  
          Fax: 319 335 9565  
          Email: msoares@blue.weeg.uiowa.edu  
          Oligo-dt track not found. Not 1 site shown in beginning of sequence  
          is likely internal to the message. cDNA library Preparation: M.B.  
          Soares Lab Clone distribution: clones will be available through  
          Research Genetics (www.resgen.com)  
          Seq primer: M13 forward  
          POLYA=No.

FEATURES  
source

1..484  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="UI-R-BJ2-bpw-f-03-0-UI"  
/clone\_1id="UI-R-BJ2"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: p773D-Pac (Pharmacia) with a modified  
polylinker. Site\_1: Not I; Site\_2: Eco RI; The UI-R-BJ2  
library is a subtracted library derived from the following  
tissues: heart, atrium at 15 dpc, ventricle at 16.5 dpc,  
atrium at 16.5 dpc, ventricle at 13 dpc, ventricle at 15  
dpc, AV canal at 15 dpc. For a detailed description of  
the library from which this clone was derived, please  
visit our web site at [ratest.eng.uiowa.edu](http://ratest.eng.uiowa.edu). The

subtraction has been previously described in (Bernaldo,  
Lennon and Soares, Genome Research 6:791-806, 1996)  
TAG\_SEQ=None found"  
BASE COUNT 98 a 137 c 164 g 85 t  
ORIGIN

Query Match 5.0%; Score 236.2; DB 148; Length 484;  
Best Local Similarity 68.2%; Pred. No. 4.5e-57;  
Matches 328; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

QY 1522 gtctcgtcccccgtatgacagaagtataaagtcaatatattgacaatctgcccctagcgg 1581  
          |||||          |||||          |||||          |||||          |||||  
Db 4 GCCCGCACCTTAGAAGAGAGGAGGACCTCGGAAACCGTACCTGGGGCTGGCCGCG 63  
          |||||          |||||          |||||          |||||          |||||  
QY 1582 attgttagaagatcaaaaagtcctgattatgcatcacagatgaacaccttaagact 1641  
          |||||          |||||          |||||          |||||          |||||  
Db 64 ATCGCAGCAAAATATTAAGTAAGTCAAGTTCGTGGCTTACAGGACGGCACCCTTTACACG 123  
          |||||          |||||          |||||          |||||          |||||  
QY 1642 cgtgaagctattcagcatgaatcaggaatcttggaccttacttatttgggaagtga 1701  
          |||||          |||||          |||||          |||||          |||||  
Db 124 COTGAGCCACCCCGCGTGAAGCTGGGCTCCTAGGCGCACTGTCTACGGGGAAGTCGA 183  
          |||||          |||||          |||||          |||||          |||||  
QY 1702 gacacactgtgatataattaaagaatcaagcaagcagacatataacatctccctcac 1761  
          |||||          |||||          |||||          |||||          |||||  
Db 184 GACTACACTGTTGATTGTGTTCAAGAACCGAGCGGCGGATACAACTCCACCTCAT 243  
          |||||          |||||          |||||          |||||          |||||  
QY 1762 ggaatcactgtctgcgccttcttattcaagagagatatacaaaaggctgaanaacattg 1821  
          |||||          |||||          |||||          |||||          |||||  
Db 244 GGAGATCCGTGACGTTGGCGCTGTGCATCAGAGAGGCTCCACAGAGGGGTGAAGACAGTG 303  
          |||||          |||||          |||||          |||||          |||||  
QY 1822 aaggaatttccaattctgcagaggaatattcaataataatgagacagtgaatga 1881  
          |||||          |||||          |||||          |||||          |||||  
Db 304 AAGGATTTGCCGATCCGCCAGAGAGACGTTCAAGTACAGTGTGAGCGTCCGACG 363  
          |||||          |||||          |||||          |||||          |||||  
QY 1882 gatgggccaactaatatcagatcctcggtgctgcgcgcgtatctcttagttcgtta 1941  
          |||||          |||||          |||||          |||||          |||||  
Db 364 GATGGGCTTGAAGGTCCGATGCGGGGTGTGACGGGCTTCAATGACGAGCGCTGTGAT 423  
          |||||          |||||          |||||          |||||          |||||  
QY 1942 atggaagaagatctagacttcagactcatgtgcccctctcattcattgtaagaatct 2001  
          |||||          |||||          |||||          |||||          |||||  
Db 424 CCCGAGAGACACTGTGGCTTCGGGCTGATTGGCCCTCTCATCTCTGTAAAGATCT 483  
          |||||          |||||          |||||          |||||          |||||  
QY 2002 g 2002  
Db 484 G 484  
RESULT 13  
BE847128 401 bp mRNA EST 26-SEP-2000  
LOCUS uw22b11.y1 Soares mouse 3NBMS Mus musculus cDNA clone IMAGE:3417405  
DEFINITION 5' similar to gb:L05573 Mus domesticus coagulation factor VIII mRNA  
          , complete (MOUSE), mRNA sequence.  
ACCESSION BE847128 GI:10305395  
VERSION BE847128.1  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 401)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
          Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
          Email: [cgaps@mail.nih.gov](mailto:cgaps@mail.nih.gov)  
          This clone is available royalty-free through INLNL; contact the  
          IMAGE Consortium ([info@image.lnl.gov](mailto:info@image.lnl.gov)) for further information.  
          MGI:1093217  
          Seq primer: -40RP from GIBCO.  
          Location/Qualifiers

## FEATURES





[illegible]

|            |  |
|------------|--|
| RESULT     | 15   |
| A1233991   |  |
| LOCUS      | A1233991 422 bp mRNA EST 08-JAN-1999                               |
| DEFINITION | EST230679 Normalized rat lung, Bento Soares Rattus sp. cDNA clone. |
| ACCESSION  | R1UCF80 3' end, mRNA sequence.                                     |
| VERSION    | A1233991 A1233991.1 GI:4135942                                     |
| KEYWORDS   | EST.   |
| SOURCE     | Rattus sp.   |
| ORGANISM   | Rattus sp.   |

|           |  |
|-----------|--|
| REFERENCE | 1 (bases 1 to 422)   |
| AUTHORS   | Lee,N.H., Glodet,A., Chandra,I., Mason,T.M., Quackenbush,J., |
| TITLE     | KerLavage,A.R. and Adams,M.D.                                |
| JOURNAL   | Rat Genome Project: Generation of a Rat EST Catalog and Rat  |
| COMMENT   | Gene Index   |
|           | Unpublished (1998)   |
|           | On Oct 30, 1998 this sequence version replaced gi:3817871.   |

The Institute for Genomic Research  
9712, Medical Center Drive, Rockville, MD 20850, USA  
Tel: (301)-838-3529  
Fax: (301)-838-0208  
Email: nhlee@tigr.org  
Seq primer: M13-21.

| FEATURES | Location/Qualifiers |
|----------|---------------------|
| source   | 1. .422             |

|            |      |       |       |      |
|------------|------|-------|-------|------|
| BASE COUNT | 74 a | 120 c | 160 g | 68 t |
| ORIGIN     |      |       |       |      |

|                           |         |                    |           |             |
|---------------------------|---------|--------------------|-----------|-------------|
| Query Match               | 4.28;   | Score 201.4;       | DB 17;    | Length 422; |
| Best Local Similarity     | -67.58; | Pred. No. 5.7e-47; |           |             |
| Matches 283; Conservative | 0;      | Mismatches 136;    | Indels 0; | Gaps 0;     |

[illegible]

| Accession | Sequence   | Position |
|-----------|--|----------|
| Dd        | 183 TGTGTATGCGCGAAGGTCTGGAGGACAGAGTGTGCTATGCTCAACATAGGGGCAC      | 242      |
| Qy        | 3542 acatccattctatcatcttcaatagtagagatgtgtcacttacaagaaaaagagagta  | 3601     |
| Dd        | 243 ACGCACAATCCGTCAATTTCACGCCGCGACACAGCTACACGCGCGAGACGGGGCGAGAC  | 302      |
| Qy        | 3602 aaatgacactgtacaactctcatccagtgcttttttagaagtgaaatgtttaccatcca | 3661     |
| Dd        | 303 GGACGGCGGTGTGCAACCTGTACCCACAGCGCTTTCAACCGTGGAGATGTTGCCGTCA   | 362      |
| Qy        | 3662 aagctgaatttggcgggttgaatggcttatttggcagacatcatagcttggatgagc   | 3720     |
| Dd        | 363 GAGCGGGATTTGGCGGGGTGAGAGACCTGGTGGCGACGACCTACCGCGCTGGGATGAGAC | 421      |

Search completed: November 18, 2001, 04:30:59  
Job time: 39941 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2001, 13:22:16 ; Search time 41.69 Seconds  
(without alignments)  
4668.282 Million cell updates/sec

Title: US-09-689-430-2  
Perfect score: 7869  
Sequence: 1 M0ELSTCFCLRLRRCFSA.....WHQIALRMEVLGCEAODLY 1471

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: SPREMBL\_16:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mhc:\*  
9: sp\_organelle:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_unclassified:\*  
13: sp\_vertebrate:\*  
14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|-------------|
| 1          | 6401   | 81.3        | 2343   | 6  | 018806      |
| 2          | 6360   | 80.8        | 2343   | 6  | 062730      |
| 3          | 2402.5 | 30.4        | 2224   | 4  | 043737      |
| 4          | 2394.5 | 30.4        | 2183   | 11 | 088783      |
| 5          | 2371.5 | 30.1        | 2258   | 6  | 096LPI      |
| 6          | 1792   | 22.8        | 1157   | 11 | 092024      |
| 7          | 1703   | 21.6        | 1084   | 11 | 09JL97      |
| 8          | 1619.5 | 20.6        | 1048   | 6  | 09X127      |
| 9          | 1334   | 17.0        | 891    | 4  | 075180      |
| 10         | 1214   | 15.4        | 782    | 4  | 075659      |
| 11         | 1104   | 14.0        | 216    | 4  | 014286      |
| 12         | 664    | 8.4         | 463    | 11 | 09R1X9      |
| 13         | 662.5  | 8.4         | 480    | 4  | 043854      |
| 14         | 661    | 8.3         | 426    | 11 | 09WTS3      |
| 15         | 655.5  | 6.5         | 480    | 11 | 035474      |
| 16         | 514.5  | 6.5         | 363    | 6  | 077718      |
| 17         | 469.5  | 6.0         | 901    | 4  | 09H2E4      |
| 18         | 469.5  | 6.0         | 901    | 4  | 09H2D5      |
| 19         | 469.5  | 6.0         | 906    | 4  | 09H2E3      |

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 20         | 469.5 | 6.0         | 906    | 4  | 09H2D4      |
| 21         | 450.5 | 5.7         | 921    | 11 | 090X38      |
| 22         | 429.5 | 5.5         | 704    | 4  | 09H2E1      |
| 23         | 429.5 | 5.5         | 779    | 4  | 09NTT3      |
| 24         | 411.5 | 5.2         | 858    | 5  | 076470      |
| 25         | 383.5 | 4.9         | 555    | 4  | 09H2E2      |
| 26         | 273   | 3.5         | 51     | 4  | 09U005      |
| 27         | 271   | 3.4         | 1128   | 11 | 088442      |
| 28         | 266   | 3.4         | 764    | 11 | 054860      |
| 29         | 262   | 3.3         | 1158   | 4  | 014113      |
| 30         | 251   | 3.2         | 3843   | 5  | 09U5D0      |
| 31         | 251   | 3.2         | 3843   | 5  | 09VU94      |
| 32         | 247   | 3.1         | 224    | 11 | 09R1M6      |
| 33         | 223   | 2.8         | 722    | 11 | 092100      |
| 34         | 223   | 2.8         | 728    | 6  | 097567      |
| 35         | 221   | 2.8         | 719    | 11 | 061281      |
| 36         | 217   | 2.8         | 734    | 4  | 09U0B5      |
| 37         | 217   | 2.8         | 1331   | 4  | 09UHC6      |
| 38         | 217   | 2.8         | 1339   | 4  | 09U012      |
| 39         | 211   | 2.7         | 192    | 11 | 090WQ1      |
| 40         | 211   | 2.7         | 1283   | 5  | 094887      |
| 41         | 211   | 2.7         | 1284   | 5  | 09VTU5      |
| 42         | 207.5 | 2.6         | 737    | 5  | 017305      |
| 43         | 206.5 | 2.6         | 737    | 5  | 018163      |
| 44         | 185.5 | 2.4         | 911    | 11 | 035407      |
| 45         | 185   | 2.4         | 3396   | 5  | 09VMS5      |

## ALIGNMENTS

RESULT 1  
018806 PRELIMINARY; PRT; 2343 AA.  
AC 018806;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE FACTOR VIII.  
GN F8.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID:9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RA Cameron C., Notley C., Hoyle S., McGlynn L., Hough C., Kamisue S.,  
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF016234; AAB87412.1; -.  
DR HSSP; P00451; ICFG.  
DR InterPro; IPR001117; -.  
DR Pfam; PF00394; Cu-oxidase; 3.  
DR Pfam; PF00754; F5\_F8\_type\_C; 2.  
DR PROSITE; PS01285; FA58C\_1; 2.  
DR PROSITE; PS01286; FA58C\_2; 2.  
DR PROSITE; PS00079; MULTICOPPER\_OXIDASE1; 3.  
DR SMART; SM00231; FA58C; 1.  
SQ SEQUENCE 2343 AA; 265829 MW; A854FME571C3B399 CRC64;

Query Match 81.3%; Score 6401; DB 6; Length 2343;  
Best Local Similarity 54.0%; Pred. No. 0;  
Matches 1268; Conservative 85; Mismatches 108; Indels 886; Gaps 4;

QY 1 M0ELSTCFCLRLRRCFSAATRRYVIGAVELSMYMSD-LGELPYARPPRPYKSPF 59  
DB 1 M0ELSTCFCLRLRRCFSAATRRYVIGAVELSMYMSDLSALHADPSFSSRVGSLPL 60  
QY 60 NTSVYKKTLFEPTVHLFNIAKRPMPMGILGPTIQAEVYDVVITLKNASHPVSLHA 119

|    |      |   |      |
|----|------|---|------|
| Db | 61   | 115T1YRTKTVEEF7DDLEFNIAKRRPPMMGLLGP7IOAQEVDYDV7IVLKNASHHVSUHA                         | 120  |
| Qy | 120  | VG5YWKASBEGAEDDOTSQRKEDDKVYPCGS3HTYVMQVLKENGPMASDPLCLTYSYLS                           | 179  |
| Db | 121  | VGVSYWKASBEGAEDDOTSQRKEDDKVYPCGS3HTYVMQVLKENGPMASDPLCLTYSYLS                          | 180  |
| Qy | 180  | HVDLIVKDLNSGLIGALLYCREGSLAKKTKOTPLHAK7ILLFAVDEKSMHSEKNSLMODR                          | 239  |
| Db | 181  | HVDLIVKDLNSGLIGALLYCREGSLAKKTKOTPLHAK7ILLFAVDEKSMHSEKNSLMODR                          | 238  |
| Qy | 240  | DAASRAMPKPKHNTGVYVNSRSLPGLIGCHRKSVYHNHVGMTTPVSHI7FLEGHTFLVRN                          | 299  |
| Db | 239  | ---AEAGHELHTINGVYVNSRSLPGLIGCHRKSVYHNHVGMTTPVSHI7FLEGHTFLVRN                          | 294  |
| Qy | 300  | HRQASLEISPTFLTAQTLMDLGOFLLEFCHISSHQHDMEAYVYKVDSCPEEBOLRMKN                            | 359  |
| Db | 295  | HRQASLEISPTFLTAQTLMDLGOFLLEFCHISSHQHDMEAYVYKVDSCPEEBOLRMKN                            | 354  |
| Qy | 360  | EEADYDODDLDSENDVVRFPDDNSPSP7IOTRSVAKKPKPKMVNHI7AAEEDMDVAPLV                           | 419  |
| Db | 355  | ED-KDYDDGLYDSMDVVSFPDDSSSP7IOTRSVAKKPKPKMVNHI7AAEEDMDVAPSGP                           | 413  |
| Qy | 420  | APDRSYKSOYLNNNGPQRIGRKKYKVRPMATYDEFKFKREALIOHSGILGP7LYGEVDT                           | 479  |
| Db | 414  | TPNDRSHKNNLYLNNGPQIGRKKYKVRVAATDEFKFKREALIOYSGILGP7LYGEVDT                            | 473  |
| Qy | 480  | LLIIFKNOASRPYNIYPG7ITDVRPDLYSRRLPKGVKHLKDP7ILPGE7IFKYKWTYVBDG                         | 539  |
| Db | 474  | LLIIFKNOASRPYNIYPG7ITDVRPDLYSRRLPKGVKHLKDP7ILPGE7IFKYKWTYVBDG                         | 533  |
| Qy | 540  | PTKSDPRLT7RYSFSSVFMERD7ASGLIGPL7LC7KESVDORGN0IMSDKRVN7LFSVD                           | 599  |
| Db | 534  | PTKSDPRLT7RYSFSSVFMERD7ASGLIGPL7LC7KESVDORGN0IMSDKRVN7LFSVD                           | 593  |
| Qy | 600  | ENRSMY7LENIORFLPNPAGVOLEPPEQOASIMHSINGVYFDS7IOL7SVC7HEVAYW7IL                         | 659  |
| Db | 594  | ENRSMY7LENIORFLPNPAGVOLEPPEQOASIMHSINGVYFDS7IOL7SVC7HEVAYW7IL                         | 653  |
| Qy | 660  | SIGATDFL7SVFSG7YF7FKHMYED7L7L7P7P7SG7E7V7MS7ENP7GL7IL7G7CHNS7P7NR                     | 719  |
| Db | 654  | SVGATDFL7SVFSG7YF7FKHMYED7L7L7P7P7SG7E7V7MS7ENP7GL7IL7G7CHNS7P7NR                     | 713  |
| Qy | 720  | GMTALLK7VSCDKN7G7DY7ED7SYED7ISAK7YL7SKNNA7IP7RFS7ONS7RNP7STR7OK7OFNAT-                | 778  |
| Db | 714  | GMTALLK7VSCDKN7G7DY7ED7SYED7IP7PL7ENN7NV7IK7P7RFS7ONS7RNP7STR7OK7OFNAT                | 773  |
| Qy | 779  | ---   | 778  |
| Db | 774  | TPENDIEK7ID7I7QSGERT7OL7IA7QSVSS7DDL7ML7LG7ONT7P7RGL7FL7SD7L7REAT7DRAD7HSR            | 833  |
| Qy | 779  | ---   | 778  |
| Db | 834  | GAIERNNK7GP7E7VAS7UR7P7EL7R7H7SE7DR7ET7PE7B7EL7OL7NL7EN7IG7T7TV7EL7KL7DL7K7ISSSS      | 893  |
| Qy | 779  | ---   | 778  |
| Db | 894  | DSLMT7SPT7IP7SDK7LAA7A7E7KT7GL7G7P7M7SV7H7NS7L7G7IV7G7N7NS7HL7I7Q7GV7PL7E7SE          | 953  |
| Qy | 779  | ---   | 778  |
| Db | 954  | EDNDSK7L7EAP7L7M7NI7Q7ESS7L7E7NV7L7S7M7ES7N7L7FK7EER7IR7G7P7AS7L7IK7ONAL7FK7VN7ISSVKT | 1013 |
| Qy | 779  | ---   | 778  |
| Db | 1014 | NRAV7VNL7TN7K7TVA7IPT7LLI7EN7ST7V7M7OD7IM7L7ERN7E7FK7E7V7SL7IN7ET7FMD7RNT7AL          | 1073 |
| Qy | 779  | ---   | 778  |
| Db | 1074 | GLNHVSNK7T7L7SK7NV7EM7A7H7OK7ED7VP7L7RA7EN7D7L7SSK7IP7FL7D7MI7K7THG7KNS7L7SEOR        | 1133 |
| Qy | 779  | ---   | 778  |
| Db | 1134 | PSPKOL7T7SL7G7SEK7V7K7DON7L7SE7EK7VV7V7GE7DE7FK7D7TEL7O7ET7FP7NKS7IF7P7AN7LAN7QEN     | 1193 |

|    |      |  |      |
|----|------|--|------|
| QY | 779  | -----  | 778  |
| Db | 1194 | DTYNOEKSPPEIEERKEKLTQENVALFOAHMTIGTKNFKNLFLSTKONVAGLEDEOYT   | 1253 |
| QY | 779  | -----  | 778  |
| Db | 1234 | PILQDTNSLNDSPHSEGIHMANSKITREENBLEGLQNOTNOWERPPTTRHSSNASOHV   | 1313 |
| QY | 779  | -----  | 778  |
| Db | 1314 | ITQNGKRSBLQPRLSQOEIKFERKVIANDTSTQMSKNMVAQGLTLOJENKEKEKRAIT   | 1373 |
| QY | 779  | -----  | 778  |
| Db | 1374 | QSPSLDCSMRNHVTIOMNDALPYAKESASPSVRHTDLTKIPSOHNSHLPASACNTYFR   | 1433 |
| QY | 779  | -----  | 778  |
| Db | 1434 | ERTSVOEGSHFLQEAKRNNLSLAFVTLGITBQGFSSLSGSATINQPMYKKEINTVLQ    | 1493 |
| QY | 779  | -----  | 778  |
| Db | 1494 | PLSLSTQKVELLSQVHVNDQEDSPRTTSNDSPGHLDLNGKIFLQTKQSPVAMKNTTSPG  | 1553 |
| QY | 779  | -----  | 778  |
| Db | 1554 | KVPFLKMTSESSEKIPSKLLGLVLANDNHDTQIPSEBKSOXKSQTNPTAFKRDTILPLG  | 1613 |
| QY | 779  | -----PVLKRHOEITRTTLQSDOEI                                    | 801  |
| Db | 1614 | PCENNDSTAALNEGODKPOREAMAKOGERGLCSQNPVSKNHOREITVTLLOPEEKEF    | 1673 |
| QY | 802  | DYDDTISYEMKKEBFDIYDEDENOSPRSFOKTRHFIYAVERLMDYGMSSPHYLARRA    | 861  |
| Db | 1674 | EYDDTFSEMKREDFDIYGEYNOGLRSFKAKTRHFIYAVERLMDYGMSSRPHILARRA    | 1733 |
| QY | 862  | QSGSVPOFKKVVFOEFTDGSFTQPLYRGELINHLGLLGYIAAEVDNIMVTFRNQASRP   | 921  |
| Db | 1734 | QSGDVQCFKKVFOEFTDGSFTQPLYRGELINHLGLLGYIAAEVDNIVVTFKNOASRP    | 1793 |
| QY | 922  | YSFSSLSLSTYEDDROGAEPRKRFVVRNENKTYFMKVQOHMAPTDDEPCKAMAFESVD   | 981  |
| Db | 1794 | YSFSSLSLSTYEDDROGAEPRKRFVVRNENKTYFMKVQOHMAPTDDEPCKAMAFESVD   | 1853 |
| QY | 982  | LEKOVHSGILPRLLYCHTNTLNAHGROVYVQEFALFTIPDETFSWYFTTEMERENCRAP  | 1041 |
| Db | 1854 | LEKOVHSGILPRLLYCHTNTLNAHGROVYVQEFALFTIPDETFSWYFTTEMLBENCARP  | 1913 |
| QY | 1042 | CNIGMEDPTEKENYRFHAINCYIMDTPLGLVMAQODRIRMYLLSMGSENINHSIHFSGHV | 1101 |
| Db | 1914 | CNVQKEDPTLKENRFHAINGYVMDTLPLGLVMAQODKVRMYLLSMGSENINHSIHFSGHV | 1973 |
| QY | 1102 | FTVYKKEEYKALYNLYPGVFETVEMILPSKAGIWRKCELIQENHLAAGSTLFLVYSNKCQ | 1161 |
| Db | 1974 | FTVYKKEEYKALYNLYPGVFETVEMILPSOVGIMRCELIQENHLAAGSTLFLVYSKCKQ  | 2033 |
| QY | 1162 | TPGLGASGHIDPQITASGOYGOMAPRLAHLHSGSINAMSKPEPFSYKVDLAPMTH      | 1221 |
| Db | 2034 | TPGLGASGHIDPQITASGOYGOMAPRLAHLHSGSINAMSKDPFSPYKVDLAPMTH      | 2093 |
| QY | 1222 | GIKTOGAROKESSLYISQFTIMYSLDQKKWQTYRGNSTGLTWVFGVNDSSGIKINIFNP  | 1281 |
| Db | 2094 | GIMTOGAROKESSLYISQFTIMYSLDQKNKHYSRGNSTGLTWVFGVNDSSGIKINIFNP  | 2153 |
| QY | 1282 | PIIARVIRLAPRTHYSIRSTLMBELGCDLNSGMPJGMEKASISDAQITASSYFTNNFAT  | 1341 |
| Db | 2154 | PIIARVIRLAPRTHYSIRSTLMBELGCDNFSCMPJGMEKASISDAQITASSYLSMAT    | 2213 |
| QY | 1342 | WSPKARLHLOGRBNARPOVNNPKEMLOYDQPKTKYLVGTVQVKSLSLMMVKEFLI      | 1401 |
| Db | 2214 | WSPQARHLHLOGRBNARPOVNNPKEMLOYDEKTKYLVGTVQVKSLSLSMVKEFLI      | 2273 |

```
QY 1402 SSSODGHWLTFEFGNGKVFVFGNODSFPTVYNSLDPLLTREYLRHPQSVWHTALRME 1461
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2274 SSSODGHWMTLPLQNGKVKVFGNNDSSPTVYNSLDPLLTREYLRHPQSVWHTALRME 2333
QY 1462 VUGCEAQ 1468
    |||||
DB 2334 VUGCDTQ 2340

RESULT 2
062730 PRELIMINARY: PRT: 2343 AA.
ID 062730:
AC 062730:
DT 01-AUG-1998 (TREMblrel. 07, Created)
DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)
DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE FACTOR VIII.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_Taxid=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY, SPLEEN;
RA Gordy P.W., Bowen R.A.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF049489; AAC05384.1; -.
DR HSBP: P0451; ICRG.
DR InterPro: IPR00421; -.
DR InterPro: IPR00117; -.
DR Pfam: PF00394; Cu-oxidase; 3.
DR Pfam: PF00754; F5_P8_type_C; 2.
DR PROSITE: PS01285; FA58C_1; 2.
DR PROSITE: PS01286; FA58C_2; 2.
DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 3.
DR SMART: SM00231; FA58C; 1.
SQ SEQUENCE 2343 AA; 265613 MW; F612D744ADADD99 CRC64;

Query Match 80.8%; Score 6360; DB 6; Length 2343;
Best Local Similarity 53.7%; Pred. No. 0;
Matches 161; Conservative 88; Mismatches 112; Indels 886; Gaps 4;

QY 1 MOEISTCFCLLRFCCSATRRYLLGAVELSDWDMQSD-IGELPYDARFPVRPKSPF 59
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 MOVELYTCCFLCLPPSLSATRKTYLGAVELSDWDMQSDLSALHADVSFSSRVGSLD 60
QY 60 NTSVYKKTLEVEFTVHLFNIAKPRPPWMLGPTIOAEVYDTVVITLKNMASHVSLHA 119
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 TTSVYRKTVFEPFDLFNIAKPRPPWMLGPTIOAEVYDTVVITLKNMASHVSLHA 120
QY 120 VGVSVKWSSEGAAYDDQTSQREKEDDKVFPGGSHYVWQVLKENGPMASDPLCTYSYS 179
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 121 VGVSVKWSSEGAAYDDQTSQREKEDDKVFPGGSHYVWQVLKENGPMASDPLCTYSYS 180
QY 180 HVDLVKDNSGLIGLALYCRGSLAKREKTOYLAHFKTLLFAVDEKSHSEKTNLMDR 239
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 181 HVDLVKDNSGLIGLALYCRGSLAKREKTOYLAHFKTLLFAVDEKSHSEKTNLMDR 238
QY 240 DAASRAVPMKHTVNGVYNSRLPGLIGLHRSKVYWHVIGMTPTPVHSLFEGHTFLVN 299
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 239 -----AEAGHEHTINGVYNSRLPGLIGLHRSKVYWHVIGMTPTPVHSLFEGHTFLVN 294
QY 300 HROASLEISPTFLTAQTLTMDLQFLFCHISSHQHDGMEAYVYKVDSCPEEPQLRMKN 359
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 295 HROASLEISPTFLTAQTLTMDLQFLFCHISSHQHDGMEAYVYKVDSCPEEPQLRMKN 354
QY 360 BEAEEDYDDLDSEMDVYRFDDNPSFQIRSVAKKHPTKVHTIAAEEEDMDAPLVL 419
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 355 ED-KYTDGLGSDMDVYSEDDSSPFIQIRSVAKKHPTKVHTIAAEEEDMDAPLVL 413
QY 420 ABDDSYKSOYLNNGPORIGRKVKVRFMAVYDEFKTRTREALIOHESGLIGPLLYGEVGT 479
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
DB 414 TPNDRSKKNLVYNNGPORIGRKVKVRFMAVYDEFKTRTREALIOHESGLIGPLLYGEVGT 473
QY 480 LLIIRKQASRPYNIYPHGINTVYRPLYSRLPKGYKHLKDFILIGELFKYKWTYVEBG 539
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 474 LLIIRKQASRPYNIYPHGINTVYRPLYSRLPKGYKHLKDFILIGELFKYKWTYVEBG 533
QY 540 PTKSDPRLTRYSSFFVNNERDLASLIGPLLCYKESVDQGNQIMSDKRWILFESVD 599
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 534 PTKSDPRLTRYSSFFVNNERDLASLIGPLLCYKESVDQGNQIMSDKRWILFESVD 593
QY 600 ENRSWYLTENIGRFLPNPAQVLEDEPFGQASNMHSHINGVPSQLSVCLHEVAYWYL 659
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 594 ENRSWYLTENIGRFLPNPAQVLEDEPFGQASNMHSHINGVPSQLSVCLHEVAYWYL 653
QY 660 SIGAOTDFLSVFSGTYTFKHKWYEDTLTFPPSGETVMSMENGLWLTGHSNDFRR 719
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 654 SVGAOTDFLSVFSGTYTFKHKWYEDTLTFPPSGETVMSMENGLWLTGHSNDFRR 713
QY 720 GMTALLKVSCKNNGDYEDSYEDISAYLLSKNNAIEPRFSQNSRHPSTROKOFNAT- 778
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 714 GMTALLKVSCKNNDYIEDYEDIPFLNENNVIKRFSQNSRHPSTROKOLKAT 773
QY 779 ----- 778
DB 774 TPENDIEKIDQSGERTOLKAQSVSSDLMLLIGNPTPRLPLSDUREATDRADDSR 833
QY 779 ----- 778
DB 834 GAIERNKGPPEVASLRPELRHSEDEFTPEPELOLRLENIGTNTVELKLDLKS 893
QY 779 ----- 778
DB 894 DSLMTSPITPSDKLAATEKTSLOPPMNSVYFNHGLGTYVGNNSHLLQSGVPLELSE 953
QY 779 ----- 778
DB 954 EDNDSKLEAPLMNIOESSLRNVLMSERNLFKEERIRGPASLJKDNALFKVNISSVT 1013
QY 779 ----- 778
DB 1014 NRAPVLTNTRKRAVAPTLLENSTSWQDIMELEERTEKFEVTSLIHNEFTMDRTAL 1073
QY 779 ----- 778
DB 1074 GLNHSNKTTLSEKNVMAHQKKEDEVPYLRAPNDLSSKIPPLPWIKTHGKNSLSSEOR 1133
QY 779 ----- 778
DB 1134 PSPKQTLISGESEKSVKQDNFLSEKVVYVGEDEFKTDTELQEIFPNKKSIFFANLANVOEN 1193
QY 779 ----- 778
DB 1194 DTYNOEKSLIEIERKEKLTQENVALPOAHWTIGTKNPLKNLFLSTKONVAGLEOPYT 1253
QY 779 ----- 778
DB 1254 PTLQDTRSLNDSPHSEGITHMANFSKIREBANLEGLGNQTNQWVERPSTTRMSSNASHV 1313
QY 779 ----- 778
DB 1314 ITQORKRSLLKQRLSOGELKFERKYIANDTSTQWSKNNMYLAQGLTLQIEVNEKERRAIT 1373
QY 779 ----- 778
DB 1374 QSPILSDCSMRNHVTIOMNDSALPVAKESAPSVRHTDYLTKIPSQHNSSHLPASACNYFER 1433
QY 779 ----- 778
DB 1434 EFTSGVQBSHPLQDAKRNNSLAEVTLGITGEGKFSLSGKSAATNOPMYKKLENTVLLQ 1493
QY 779 ----- 778
DB 1494 POLSETSDKVELLSQVNHQDEDSFPTKTSNDSRGLHLDLKGITFLQKTGQPVKMNKNTNSPG 1553
```

QY 779 ----- 778  
 Db 1554 KVPFLKWTESSEKIPSKLLGLVLA MDNHVDTQIPSEEMKSOQKSTNTAFKRPDLPLG 1613  
 QY 779 ----- PVLKRNORITITTTLODOEEI 801  
 Db 1614 PCENNDSTAINEGODKPOREAMMAKOGEPRLCSONPVPKSHOREITVTTLQPEEDKF 1673  
 QY 802 DVEDDTISVEMKKEPDIDYDEDENOQSPRSFOKTRHYFIAVERLMDVGMSSPHVLNRA 861  
 Db 1674 EYDDTFSIEKRRDPDITGDEYDQGLRSFOKTRHYFIAVERLMDVGMSSPHVLNRA 1733  
 QY 862 OSGSVPOFKKVVYQFETDSEFTOPLYRGELNEHLGLLPYIRA EVEDNIMVTERNOASRP 921  
 Db 1734 OSGDVQOFAKKVYQFETDSEFTOPLYRGELNEHLGLLPYIRA EVEDNIMVTERNOASRP 1793  
 QY 922 YSFYSSLLIYEDOROGAERPRKRVFNPKTKTFKVKYQHMAATPKDEFDCKANAYFSDVD 981  
 Db 1794 YSFYSSLLIYDEDEGGAERPRKRVFNPKTKTFKVKYQHMAATPKDEFDCKANAYFSDVD 1853  
 QY 982 LEKDVHSGILGPLVCHTNTLPAHGRQVTVQEFALFETTFDETKSVYPTENNERCRAP 1041  
 Db 1854 LEKDVHSGILGPLVCHTNTLPAHGRQVTVQEFALFETTFDETKSVYPTENNERCRAP 1913  
 QY 1042 CNIQMEDPTFKENYRPHAINGYIMDTPLGLVMAODQIRMYLLSMGSNENIHSTHESGHV 1101  
 Db 1914 CNVQKEDPTLKEFRPHAINGYIMDTPLGLVMAODQIRMYLLSMGSNENIHSTHESGHV 1973  
 QY 1102 FVVRKKEEKMAKLVNLYPGVFETVEMLPKSGIWRVBCILGELHAGMSTLFLVYSKQCQ 1161  
 Db 1974 FVVRKKEEKMAKLVNLYPGVFETVEMLPKSGIWRVBCILGELHAGMSTLFLVYSKQCQ 2033  
 QY 1162 TPLGMAAGHIDPOLTASGOYOMAPKLARLHSGSINAMSTKEPFSMIYVDLAPMTH 1221  
 Db 2034 TPLGMAAGHIDPOLTASGOYOMAPKLARLHSGSINAMSTKEPFSMIYVDLAPMTH 2093  
 QY 1222 GIKTGOAROKFSSLYISOFTIMYSLDGKKWQYRGNSGTGLMWFCGNVDSGKIHNIENP 1281  
 Db 2094 GIMTGOAROKFSSLYISOFTIMYSLDGKKWQYRGNSGTGLMWFCGNVDSGKIHNIENP 2153  
 QY 1282 PIIARIYIRLHPHYSTRSLRMELMGC DLNCSMPLEGMEKASIDAOITASSYTFNMFEAT 1341  
 Db 2154 PIIARIYIRLHPHYSTRSLRMELMGC DLNCSMPLEGMEKASIDAOITASSYTFNMFEAT 2213  
 QY 1342 WSPSKARLHLOGRSNAMPQOVNPKEMLOYDFQKTKKVTGVTTOGVKSLTSMYVKEFLI 1401  
 Db 2214 WSPSKARLHLOGRSNAMPQOVNPKEMLOYDFQKTKKVTGVTTOGVKSLTSMYVKEFLI 2273  
 QY 1402 SSSQDGHQWTFEFGNGKRVKVFQGNODSFTRPVNSLDPELLTRYLRILHPQSVHQAIALRME 1461  
 Db 2274 SSSQDGHQWTFEFGNGKRVKVFQGNODSFTRPVNSLDPELLTRYLRILHPQSVHQAIALRME 2333  
 QY 1462 VLGCEAQ 1468  
 Db 2334 VLGCDTQ 2340

RESULT 3  
 043737 PRELIMINARY: PRT: 2224 AA.  
 AC 043737  
 DT 01-JUN-1998 (Tremblrel. 06, Created)  
 DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)  
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
 DE FACTOR V.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NCBI\_Taxid=9606;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RA Bird C.;

RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: 299572; CAB16748.1; -  
 DR HSSP: P00450; 1KCW.  
 DR InterPro: IPR000421; -  
 DR InterPro: IPR001117; -  
 DR Pfam: PF00394; Cu-oxidase; 3.  
 DR Pfam: PF00754; F5\_F8\_type\_C; 2.  
 DR PROSITE: PS01285; FA58C\_1; 2.  
 DR PROSITE: PS01286; FA58C\_2; 2.  
 DR PROSITE: PS00079; MULTICOPPER\_OXIDASE1; 2.  
 DR SMART: SM00231; FA58C; 1.  
 KW Blood coagulation.  
 SQ SEQUENCE 2224 AA; 251673 MW; 1013108049690EAB CRC64;

Query Match 30.5%; Score 2402.5; DB 4; Length 2224;  
 Best Local Similarity 26.3%; Pred. No. 8 8e-177;  
 Matches 598; Conservative 280; Mismatches 489; Indels 909; Gaps 37;

QY 22 RRYVLGAVELSDWYMOISDLGELPVDARFPRVPSPPFTSVV-YKRTLVEFTVHLFNI 80  
 Db 32 RQFYVAAQGISWSVRPE-----PTNSLNLSTVSFKKIYYREVEPY-FKK 75  
 QY 81 AKRPPPMGLGPTIOAEVYDVTYITLKMAHSIPVGLHAVGVSYKASGAEYDQTSOR 140  
 Db 76 EKPOSTISGLGPTLYAEVGDIIKVEKNKADKPLSIHQGLRYSKLSGASAYLDHTPPA 135  
 QY 141 EKEDKVEFPGSHTYWQVLEKENGPMASDPLCTYLSYSHVDLVKLNGLGALLVCE 200  
 Db 136 EKMDDAVAPRETYETKESISEDGPTHDPPCLTHIYSHENILIEFNSGLGLPILICK 195  
 QY 201 GSLAKEKQ-TLHK-FLLEFAVEDEKSWHSETKNSLMODRAASARAMPKHTVNGYVN 258  
 Db 196 GTLLEGCTQTFPKQVLYLFAVFEDEKSMQSSS-----LMTYGVVN 239  
 QY 259 RSLPGLIGCRKSVYVWVAVIGMTPEVHSIFLEGHFLVRNHRQASLESPITFLAQT 318  
 Db 240 GTPMDITVCAHDIISWHLGLMSGPELFSIFNGQVLEONHHRVSAITLVASTTANNT 299  
 QY 319 IMLDGOFLCHSHSHOHGMEAYRVKVDSCPEPQLRMKNNFEADYDDDLIDSEMDYVR 378  
 Db 300 VPEEGKWIISLTPKHLQAGMCAVYIDIKKCPKTRMLKITYE----- 342  
 QY 379 FDDNDSPTQIRSVAKKHKRTVWVYIAAEEEDMDYAPLVLPDDRYSKYQYLNNGPQRI 438  
 Db 343 -----QRRHKRWVEYFLAAEVIWMDYAPVTPANMDKKYKSHLDNFSNQI 387  
 QY 439 GKRYKVRPMATYDETFKTRREAL---QHESGILGPLLYEVDITLITKNOASRPYNI 495  
 Db 388 GKHYKKVMYQYDEDSF-TKHVTPNPKMKGDIIGPIIRAQVBDTLKYKKNASRPYSIY 446  
 QY 496 PHGIT-----DVRPLYSRLPKGVKHLKDFILPGLIEIKYKMTVTEDEPKSPDRCLT 549  
 Db 447 PHGVTFSPYDEVNSFT-----SGRNNTMIRAVQPEBTYTKMNLIEPPEPENOACLT 502  
 QY 550 RYYSSEFVNMERDIASGLIGPLICYKESVDQROGQIMSDKRVNLLFSFVENDRSWLTEN 609  
 Db 503 RPYSDVDIMRDIASGLIGLILICKRSRDIQIRAADIEQAAVFAVAVDEKSKWYLEDN 562  
 QY 610 IQRFLEPNPAGVULEDPEFOASNTMHSINGVYPSL-QLSVCLHEVAWYILSLIGQDTL 668  
 Db 563 INKFCENPDEVKRDDRPFESNTIMSTINGVPSITTLGFCDDTVOMHFCSVGTONEL 622  
 QY 669 SVFSGYTFKHKVAVYDITLLEFSGEYFVMSNENGLILICHSN----- 714  
 Db 623 TINFTHGSLFYGRKHEDITLTFPMRGEVYTIMDNVGTMLLSNMSSPRSKRLKFPDY 682  
 QY 715 -----DFRNR----- 719  
 Db 683 KCIIPDDEDSYEIFEPPESTVMAIRKMDRLPEDEDESDADYDQVRLAALGIRSFNS 742  
 QY 720 -----GMTAL----- 724

```
Db 743 SLNDEEEFNLTALENGTEFEVSSNTDIIIVGSNTSSPSNISKFTVNNIAEPQAKPSHQO 802
QY 725 -----LKVSSCDKNTGDIYEDSYED----- 744
Db 803 ATTAGSPLRLHIGKNSVNSSTAHSSEPYSEDPIEDLPQDVGTGRLSLGAGEKFSOE 862
QY 745 ----- 744
Db 863 AKHKGPKVERDQAAKHREPMWKLAAHKVGRHLSQDTGSPSGMRWEDLPQDGTGSPSMR 922
QY 745 -----ISAYLSKNN----- 755
Db 923 PWDKPPDLLLKOSNSSKILVGRWHLASEKSGYEIIODTDEDTPAVNNMLISQNASRAW 982
QY 756 ----- 755
Db 983 GESPPLANKPGKQGHKFPVRVRRKSLQVRODGGKSLKKQFLKTKRKKKEKHTHAP 1042
QY 756 IEPNSF-----SQ 763
Db 1043 LSPRFHPLRSEAYNTFSERLKLHSLVAKSNTSLPTDLNQTLPMDFGMTASLPDHNQ 1102
QY 764 NSRH-----PSTR----- 771
Db 1103 NSSNDTGACSPGLYQTVPEEHYQTFPIQDPDMHSTSDPSHSSSPELSEMLEYDRS 1162
QY 772 OKOF-----NATPPYLKRH- 785
Db 1163 HKSPFTDISQSPSSSEHEVWQTVISPDLSQVTLSPELSQTNLSPDLSHHTTLPRLQNL 1222
QY 786 ----- 785
Db 1223 SPALGOMPISPDLSHTTLPDLSHHTTSLDLISQTNLSPELSQTNLSPALGOMPLSPDL 1282
QY 786 ----- 785
Db 1283 TTLSLDSQTNLSPELSHMTLSPELSQTNLSPALGOMPISPDLSHTTSLDPSQTNLSPE 1342
QY 786 -----QREITRTTLOSQDEI----- 801
Db 1343 LSQTNLSPALGOMPLSPDPSHTTSLDLISQTNLSPELSQTNLSPDLSEMPPLADLSQPL 1402
QY 802 ----- 801
Db 1403 TPDLQMTLSPDLGDTLSPNFGMSLSPDLISQVTLSPDISDTTLPLDLSISPPDLQ 1462
QY 802 -----DYDDT 806
Db 1463 IFYSESSOSLLOEFNESPYPDLGOMPSPSPPTLNDTFLSKEPNPLVIGLSKGDHY 1522
QY 807 ISVEKKF-----DPDIYDE-----DENOSP-----RSFOKKTTHY 837
Db 1523 IEIIPKEVQSSSEDYAEIDVYPDDPKTVDRTNINSSRDPDITAAWYLLKSNNGNRNY 1582
QY 838 FIAAVERLMDYGSSSPVHLNRR--AOSGSVPO--FKVVFQEFQDGSFQPLYRGELN 892
Db 1583 YIAAEISMDV-----SERVQRETDIEDSDDIPEDTYYKVVFRKYLDDSTFKRDPRGYE 1638
QY 893 EHLGLGFIYAAVEDNIMTFRNQASRPYSFYSSLSYE-----EDROGAEPKRN 944
Db 1639 EHLGLGFIYAAVEDNIVQVFRKYNLASRPYSLHAHGLSEKSSCKGTEDSDPEWFKEDN 1698
QY 945 FVKRNETFTYFMKVOHNAAPRKDEFDCAMAYVFGDVLLEKYNHSGLLGILLVCHTNLNP 1004
Db 1699 AVQPSSTTYVWHATENSGRPSGACRAMAYVSAVNPCKIHSGLLGLPLLCOKGILHK 1758
QY 1005 AHGROVTOEALFEFTJFDETKSWFTENMERNCRAPCNIOMEDPTFENYRFAINNGY 1064
Db 1759 DSNMPMQMRREVLLFMTDEKKSWMYEKSSRSMR-----LTSSEKKSHFEAHINMI 1812
QY 1065 MDTLPLGVLMAODOKIRMYLLSMGSNENITHSGHVFYVKKKEEYKALYNLYPGVET 1124
Db 1813 Y-SLPGLKMYEQEWVRLHLLNIGGSODIHVVFHGQTLLENGNKQHOLGVMPPLPGSPKT 1871
```

```
QY 1125 VEMLEPKAGIRWECLIGEHLHAGCSTLFLVYSNKCOTPLGMAQHIDPQITASGGYQ 1184
Db 1872 LEMASKRGWMLLNTVEENORAGCQTFPLIMDRCPHMLSTGIIISDSIKKSEFLGY 1931
QY 1185 WAPKLARLYSGSINAWSTKE--PFS--WIKYDLAPMTIHGIKTQAGAROKFSSLYIS 1238
Db 1932 WEPLRLANGSGSYNAWSEVKLAEPASKPWIQVDMQKEVLIITGIGTQGAHYLKSCYTT 1991
QY 1239 QFIIMYSLDGKKMQTYKRNSTGLMVFEGVNDSSGIRKININPPLIAXYIRLHPHYSIR 1298
Db 1992 EFYVAYSSNQINMOIFKGNSTRNRYFNGNSDASYIKENQDPPIVARYIRISPRAYNR 2051
QY 1299 STLNMELMGCDLNSCMGESHKASIDQITASSYFNFMAT--WSPSKARLHLOGRGA 1357
Db 2052 PTLRLLEOGCEYNGCSTPLGEMNGKIEKKOITASSFKKSMGWDEPRARLNMQRVNA 2111
QY 1358 WRPOVNNPKEMLOVDFQTKMVTGTTQGVKSLTSMYKBEFLISSQDGHQWTLFFQNG 1417
Db 2112 WQAKANNKKQWLEIDLKIKKITAITTGCKSLSEMYVKSQYTHYSQGVEMKPRYLS 2171
QY 1418 KV--KVFOGNDSTFPVYNSLDPPLTRYLRHPOSWVHOTALREYLGCEADOLY 1471
Db 2172 SMVDKIFEGNNTKGVHKNFNPPIISRIFRIVPXTMOSIALRLLEIFGC---DIY 2224

RESULT 4
088783 PRELIMINARY; PRT: 2183 AA.
AC 088783;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE MURINE COAGULATION FACTOR V.
GN F5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang T.L., Cui J., Rehumtulla A., Mousallil M., Kaufman R.J.,
RA Glosburg D.;
RT "The structure and function of murine factor V and its inactivation by
RT protein C.";
RL Blood 91:0-0(0012).
RN [2]
RP SEQUENCE FROM N.A.
RA Glosburg D., Yang T.L., Cui J., Yang A.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U52925; AAC93953.1; -.
DR HSSP; P00450; IKCW.
DR MGD; MGI:88382; F5.
DR InterPro; IPR000421; -.
DR InterPro; IPR001117; -.
DR Pfam; PF00394; Cu-oxidase; 3.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 2.
DR SMART; SM00231; FA58C; 1.
SQ SEQUENCE 2183 AA: 247228 MW: BFOA8AA723F60317 CRC64;

Query Match 30.4%; Score 2394.5; DB 11; Length 2183;
Best Local Similarity 26.0%; Pred. No. 3.6e-176;
Matches 588; Conservative 296; Mismatches 494; Indels 885; Gaps 35;

QY 8 CFFLCCL--KFC-----FSATRRYLGAVELSDMYMSDGLPYDARFPFRPKS 56
Db 7 CFFLLVYLGTGVMAGWGHQAQALRFQYVAAGQILWNYHDE-----PTD 51
QY 57 PPFMTSVYVYKKTLEVEFTVHLFNIAKPPRPWMGILGPTIQAEVYDVTVITLKNMASHPV 116
```



Db 52 PSLNIPSEFKKIVREYEOY-FKKEKPRSSNSGLGPTLYAEVDGIKHFNRKAKRPLS 110  
QY 117 LHAVSVYWKASGEAYDQOTSOREKEDKVPFGSGSHTYVMOLKENGMAAPLCLTYS 176  
Db 111 IHPGKIKTSKSEBASADHTFPAERKDOAVARGEYETTEMWYSEDSGPTPDDPLTLH 170  
QY 177 YLSHVDLVKDNSGLIGALLVCREGSLAKEKTOTL--HKFILLFAVDEGKSHSETKNS 234  
Db 171 YYSYENLQDFENSGILGPLLICKKGTLEDGTQKMFQKHVLLFAVDESKS----- 222  
QY 235 LMODDASARAKPMKMHYNGVNSPLGLICHRKSVWHYIGMTPEVHSIFLEGHT 294  
Db 223 -----RSOSPMLTYTNGFNKTMPTDITVCAHDVSWHLGMSGSELPSIHNGOV 274  
QY 295 FLVRNHRQASLEISPTFLTAOTLLMDLQGLFCHISSHQHDMGAAVYKVDSPREPOL 354  
Db 275 LEQONHAKSYTVLVSATSTTAMTMSPEGRWYSSILPRHYQAGMAYIDIKMCPKTR- 333  
QY 355 RKMKNNEAEDYDDDLTDEMDVYRFDDDNSPSFTQIRSAVKHHPKTMVHYIAAEEEDMDY 414  
Db 334 -----SPK--TLTREQRRYMKRMEYFIAAEVITMY 362  
QY 415 APVLPADPRSKSQYLNNGPQIRGKRYKRYRMAVTDFT--KTREAIQHSGLIGPL 472  
Db 363 APVLPANDKTIYRSQHLNDFSNOIGKHYKYIYROYEETFTKTDNPSIKSGILGPVY 422  
QY 473 YGEVDDTLIFIKNOASRPYNIYPHGITDVRPLYSRLKGY-----KHLKDEPILG 525  
Db 423 RAQVADTLKIYFKNPASRYSIYPHGV-----FSPYEDGINSSTSGSHTTIRVOYG 476  
QY 526 EIKFKWTVVEDGPTKSDPRCLTRYSSFVNMERDLASGLIGPLICRYESVDQGNQI 585  
Db 477 EFTYKWMILEREDEPTENDAOCLTRPYSDVYTRDIAAGLIGLILLCRSRLDQGVOR 536  
QY 586 MSDRNVILFSYFDENRSMYLTENIQRLPNPAGVLEPPEOASIMHNSINGYVDSIQ 645  
Db 537 VADLEQAVFAVFEENKSMYIEDINIKFENDEVKRDPKEYESIMSTINGYVESIS 596  
QY 646 -LSVCLHVAWYILSIGAOTDFLSVFSGYFHKHMYVEDTLTPPFCGYFMSMEMP 704  
Db 597 TLGCFDFTYQMHFCSYVGHHDILITIHFGHSFYIGRRHEDTLTLPFMGESYTYMDVY 656  
QY 705 GLWILGCHNSDFRNGMTALLKVSSCDKNTGY-YEDSYE----- 743  
Db 657 GTWMLTJNNSMFKRNLRLRFRDYKCNR--DYDNEDSYEIEYEPAPTSMTRIRIDSLE 713  
QY 744 -----DISAYLLS-----KNAIEPR----- 759  
Db 714 NEFGIDNEDDYOYLASSLAGIRSFKNSSLNPEENEFNLALALENSEFISPSTDVVD 773  
QY 760 ----- 759  
Db 774 SNSSRLSKIIINNKLQFORLTPGSGAIYAGTLNLNLGLDENFVLSNSTERSSSYHEN 833  
QY 760 ----- 759  
Db 834 DMENPQSNITWYLLPLGPKSGNREODKPKTKTGRPHMKHRSWMKAPAGKTCGRHSN 893  
QY 760 ----- 759  
Db 894 PKNSYSGMKSEEDIPSELLPKOKITSKFLNRMRVASEKGYETIAANGEDTDVDKLN 953  
QY 760 -----SFSQNSRHP-----TRO 772  
Db 954 SPQONITVPAGESISHNTTRKPPSDLPTFGSGHKSHPVROEBENSQOKROLFLRTKK 1013  
QY 773 KO----- 774  
Db 1014 KKKNNKLLALHSPRGPDLGHNHSPRPDRRLNLHSLHKSMTALSPDLNQTSPSM 1073  
QY 775 -----FNATRP 780  
:

Db 1074 STDRSLPDYNOYSKNDTEQMSSSLDYOSVPAEHSPTFPADPDQHTSTDPYRSSPP 1133  
QY 781 VLKR----- 784  
Db 1134 ELSQGLDYDLSDHFPDDIGLTSFPFQSQKSSPSDDQALPSSDLSLFTISPDLQTI 1193  
QY 785 ----- 784  
Db 1194 IYPDLDQLLSPEDNQKTSPPDLGVPLSPDDNQKTSPPDLQVLSLSPDDNQKTSPPDLG 1253  
QY 785 -----HOR----- 787  
Db 1254 QVPLSLDDNQKTSPPDLGVPLSPDDNQMTSPDLGVPLSSDDNQKTSPPDLGVPLFPE 1313  
QY 788 -----EITRTTLOSDOE----- 800  
Db 1314 DNQNTYLDLSQVPLSSDQNETSSTDLTLSPDFQOTYVSPDLQPLPSDQVTVSPD 1373  
QY 801 ----- 800  
Db 1374 LSLTLSPDFNELIAPDLGVTLSPDLIQTNPALNHGKASADPDQASYPDSQASS 1433  
QY 801 -----IDYDDT----- 806  
Db 1434 LPENLTLPDLTLHPSPSPSLNTNLSLRKNPLVYVGLSRVGDVBEIVPSEPER 1493  
QY 807 ISVEMKKEDEFDYDEDENSPSPFOKTRH-----YFLAAVERLMDY 848  
Db 1494 IDEDYAEDEFYVYNDYRFDITDYVNSSRNPDTLAWLIRGHGKAKRYTYIAAEITWY 1553  
QY 849 GMSSEPHVLRNRAQSGVPO--FKVYVQFETDGSFTQPLVRGBLNLHLLPLPIRAE 905  
Db 1554 AEPQASEM--DHEDGHTPKOTTYKKVYRYKLYDSTFSRDBRAVEEHLGILGPVIRAE 1611  
QY 906 VEDNIMWTRNRAASRPYSYSSLSYE-----EDORQAEPRKNVKNENEKTYRWK 957  
Db 1612 VDDVIOVRKNLASRPYSLSHAGLSYESSSEKTYEDESPEWFODDAVQONSSYTYWH 1671  
QY 958 VOHHNAPTDEDFCKAMAYFSDVBLEKDVHSGLIGPLLVCHNTNLPNAGROVYQOEFAL 1017  
Db 1672 ATRKSRPEMRPGSACRMAAYSAVNERDHSGLITPRLICRGTIHAMRNLPMKREPLV 1731  
QY 1018 FETIPEDEKSWYFTEEMERNRACPNIQMEDPTEKENTRFAHINCYIMDTLPGLVMAODQ 1077  
Db 1732 LPMVFDEKSKSWYERS--KGSR-----RIESPEEKNAKFYALINQMIVN-LPGRLMYEOE 1783  
QY 1078 RIRWYLLMSGNENHSHIFSGHVFYVRKKEEYKALNVLPGVEYTEMLEPSKAGIWRV 1137  
Db 1784 WYRLHLNLNNGSRDLHVHFGOTLLDNKTOHQLGVWPLRPGSFKTELEMAKSRGWWML 1843  
QY 1138 ECLIGENLHAGSTLFLVYSNKCQTPRLGMAAGHIDPQITASGOYQWAPKLARLHYSGS 1197  
Db 1844 DTEVGNQVAGMOQTPRLIIDLKCKMPMGLSTGIVISDSQIKASEYITLYPERLARLNNAQS 1903  
QY 1198 INAWSTKE-----PPS-WIKVDLAPMITHGIKTOGAROKRESSLYISOFITMYSLDGCKW 1251  
Db 1904 YNASTIEKTALDPFIPKIQVDMQKEVVYTGIOGAKHYLKSCTPTEFOVAYSSDQNMW 1963  
QY 1252 QYIRGNSGTIAWVFRGNVSSGIKHNIFNPRTIAYIRLHPHYHSIRSTIKRELMLGCDLN 1311  
Db 1964 QIFRGSKGSVWYFTGNSGSGSTIKENRDLPRIVARYIRHPHYHSYNRPLRLTLOGGCEVN 2023  
QY 1312 SCSEMPLGESKALISDAQITASSYFTNMPAT--WSPSKARLHLOGRSNARPOVNNPKEWLO 1370  
Db 2024 GCSITPLGEDGRIOKQOITASSFKKSMGDIYERPEARLARNQGRVNAQAANNKKNQWLO 2093  
QY 1371 VDFQKTMKYVTGTOGVKSLTSMYVKEFLISSQDGHQWTLTFQNGKV--KVFQGNQDS 1428  
Db 2084 VDLKIKKYTAIVTQCKSLSEMYKSYIOYSDQVAMKPYROKSSMYVCKEIGNSNT 2143  
QY 1429 FTRPVANSIDPRLTXYLRKHQSWYHQTALRMEVLEGCAQDLY 1471  
Db 2144 KGHMKNFNPPLISRFRIRIIPKTMQSIARLELFGC---DIY 2183

RESULT 5  
 O9GLP1 PRELIMINARY; PRT: 2258 AA.  
 ID O9GLP1  
 AC O9GLP1;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE COAGULATION FACTOR V.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RA Grimm D.R., Colter M.B., Kim H.;  
 RT "Cloning of the complete cDNA sequences encoding porcine factor V and  
 protein C";  
 RL Submitted (Oct-1999) to the EMBL/Genbank/DBD databases.  
 DR EMBL: AF191308; AAC28381.1;  
 SQ SEQUENCE 2258 AA; 256078 MW; 9159B9E0076A2ACC CRC64;

Query Match 30.1%; Score 2371.5; DB 6; Length 2258;  
 Best Local Similarity 25.7%; Pred No. 2,3e-174;

Matches 594; Conservative 277; Mismatches 493; Indels 949; Gaps 35;

QY 22 RRYLIGAVELSDYQMSDGLGLPVDAARPPRPVPSFPTSVYKKTLFVEFTVHLENIA 81  
 Db 32 RQYVAAGSISNMVHPE-----PTHSSSPFATS--FKIIVREYAY-FQKE 76  
 QY 82 KPRPPMGLLPTIQAEVYDTVYITLKMAHPSVSLHANGVSYWKASEAEDDQTSQRE 141  
 Db 77 KPRSPMSGLLPTIADVDIMKVHFRNKADKPLSIHPGIKYKFAEGASYPDHFTFLVE 136  
 QY 142 KEDDKVPGSGSHYVQVQLKENGPMASDPLCLTYSLSHYVDLYKDNLSLIGALLGALCREG 201  
 Db 137 KMDAAAPGQGYEYEMNISDSGTHNDPCLHIIYSENLIDQNSGLIGLICKRG 196  
 QY 202 SLAKETQTL--HKFLFLRAVEDEGKSMHSEETKNSIMQDRDAASARAMPKMTGVNVR 259  
 Db 197 LTLEDIOKMPKQYVLMRAVEDEKSNQSS-----LMTVNGVYNG 240  
 QY 260 SLPLGLIKHRKSYVHVIGMGTTPREYHSIFLBSGHTFLVNNHROASLEIPITFLAQTLL 319  
 Db 241 TMDPIVCAVDHISWHLIGMSSGPELFSIHFGQVLEONHKKASATLVASATSTANMTV 300  
 QY 320 MDLQGLFLCHLISHHODGMEAVVVKDSCPEEPQLMKNNNEAEVDDDLTPSEKMYVRF 379  
 Db 301 SPEGKMPISLLPKHFQAGQAIIDIKCAKTKRPPK-----LTRDQ----- 343  
 QY 380 DDNSPSFIOISVAKKHRTWVHYIAAEEEDDYAPLVLPADRSYKSOYLNNQFORIG 439  
 Db 344 -----RRHKRWMEYFIAAEEVIMDYAPILIPANMDKRYRSLHLDNFENQIG 388  
 QY 440 RYKTKRYFMAVYDETFKTR--EAIQHESSILGFLYGEVGDILLIFKQASRPINIIYH 497  
 Db 389 KHKKKVYKQYODESTFKLENNPNKKEDGILGPVIRAOVDELKIFKMMASRSYSIYH 448  
 QY 498 GIT-----DVRLYRRLPKGYKHLKDRPILPGLIEFKYKMYTVYDGGTSPRGLTRY 551  
 Db 449 GYTFSEYEDDVSSSTSDNNMTIR-----AVDGEYTYTKMNLDESDEPTENDAOCLTRP 503  
 QY 552 YSFFVMNERDLASGLIGPLLCYKESVDORGQDMSDKRNVLISYFDENRSGYULENTIO 611  
 Db 504 YSVDVITRDIASGLIGLLICKSRSLDRGIORTADIQKAVNFVFNDEKMKMYIEDNLY 563  
 QY 612 RELPNAGVQLEDPEFOASNIHMSINGVYFDSL-QLSVCLHEAVAYLISIGACDFELSV 670  
 Db 564 KCEMPEKVRDDPKPEYESNIMSTINGVPEPIPLTGFCFDDTVQMHFCSVTRHDNILLTI 623

QY 671 FFSGYTFKKHKNVYEDTLTLFPGSETFVMSMNPGLMILGCHNSDFRNR----- 719  
 Db 624 HFTGHSFIYKGRKHEDTLTLFPMRGESVYTMQVGMVMTLTNNSPNRKKLQKTRDYKC 683  
 QY 720 ----- 719  
 Db 684 IRDDDEDSYEIIYEPSSSTLTATRKMHDSSENKEENDEYDYQDLLASVLGIRSFRRSS 743  
 QY 720 -----GMTAL----- 724  
 Db 744 LYQEDDEFMLTALLENSEFIPTSTDAVDSSSPGNSIRAPANTFTEPRKILPHRPA 803  
 QY 725 LVYSSCDKMTG-----DYED----- 740  
 Db 804 TKAGSPRRHTGLVKNLVLRKRTQHSDPYSEDP IENPLOSITGSLDPFGTEGFRNRKH 863  
 QY 741 -----SYEDISAYLL----- 750  
 Db 864 PKHKRPKAGRDQAKHRFSQMEFPAHKGTGRHISQDNSSSSSKMGPLEDLSDDLLEKRPD 923  
 QY 751 -----SKNA----- 755  
 Db 924 STINGKMLVSEKSYEIVQADADEMAVKNLPPNNQNASRSGENIPTNKHGKORGHP 983  
 QY 756 -----IEPRFS----- 762  
 Db 984 FYTRAKHLODEODEGNSILKKGFRPIRTKRKKKERPVHNVPLSPRSFPLRGEANTPRS 1043  
 QY 763 ---QN-----SNHPRQKQFNAT----- 778  
 Db 1044 DRQNHSLLSHNSNETFPRTLNQTPPSMNLISLASHPDHQNLPMDTHQTSPLDYQT 1103  
 QY 779 -----PVYLRHQKEIT----- 790  
 Db 1104 VTPDEYQTAPIODDLPTSTAVPSHOSSLPEIQMDHDLRNKASPTDVSMEFSLKX 1163  
 QY 791 ----- 790  
 Db 1164 AGRHTTSPDLNQTSLSPELSQTTLSPDPGCHVTLSPDLQTTTLSPDLGHTTL 1223  
 QY 791 ----- 790  
 Db 1224 SPDLSTTLSPDLQTTTLSPDLSTHTTLSPDLGHTTLSPDLGHTTLSPDLGHTTLSPDL 1283  
 QY 791 ----- 790  
 Db 1284 TPLSPDLGHTTLSPDLGHTTLSPDLSPDLGHTTLSPDLGHTTLSPDLGHTTLSPDL 1343  
 QY 791 ----- 790  
 Db 1344 LSHHTLSPDLGHTTLSPDLQTTTLSPDLGHTTLSPDLGHTTLSPDLGHTTLSPDL 1403  
 QY 791 ----- 790  
 Db 1404 SPDLGHTTLSPDLGHTTLSPDLSPDLGHTTLSPDLGHTTLSPDLGHTTLSPDLGHTTL 1463  
 QY 791 -----RTIQSD----- 797  
 Db 1464 IFFSPDLQTTTLSPDLNETTLSPDLKQTSPPHDPKTSYISSESQSVTLPERGQTSPPPD 1523  
 QY 798 -----OBEIDYDQTSVEMK 813  
 Db 1524 LGQRPSPSHSLNMTFLPREFNPAVVUGLSKDDQDYVEIIPRQGEENSEEDYVXIDIVE 1583  
 QY 814 EDFDIYDED-----ENOSP-----RSFOKTRHYFIAAVERLYMDYGMSSSPHYLRN 859  
 Db 1584 YD-DYQTDVTRDINSRPNQIYAMYLRSNNGNRNYIAAELSMOYSK-----FTQ 1636  
 QY 860 RAQSGSVPO---FKVYVFOETDGSFTORLYRGELEHILGLIPYIRAEVEDNINWTFRN 916  
 Db 1637 REDIDVPEHTIYKRVVFKYDSTFTKLDPRGEVEENHIGIGIPIRAEVDDVIOVRFXN 1696  
 QY 917 QASRPYSFSSLSYE-----EDQRGAGPRKRNPFVKNPTKYTFPKVQVGNMARTKDE 968

```
Db 1697 LMSRPSLAHAGLSYKSSBSGKTYEDDSPEMFEDNAVQNSSTYVWMAHTEKSPSPG 1756
Oy 969 FQCKAAVFSVDVLEKDVHSLGLPLLVNTLNPAHQVQVQOEALFETIDEPKSW 1028
Db 1757 SACRAMAYTSAVPEKDIHSGILGLPLCKGTLHKENMPDREVVLPMVFDEKSW 1816
Oy 1029 YFTENNERNCRAPCNIOMEDPTEKREYRFAHNGIYIMDLPLGLVMAODRIHWLTLMSGS 1088
Db 1817 YXEKFTRSWR-----LTSEVKNSHFKHAINGMAYN-LPGLMAYQOEVMRLHLMLGG 1869
Oy 1089 NENIHSIHSGHVFYRKKEEYKALNLYPGVFETEMLPKAGITRVRECLIGEHLHAG 1148
Db 1870 SBDIHVHGHGOTLLENGTOHQHGWPLPLPGSEFKTELEMTKSAGMWLLDTEVEENORAG 1929
Oy 1149 MSTLFLVYNSKQTPGLGASGHIRDFOITASGOYGQWAPKARLHLYSGSINAMTEKPEFS 1208
Db 1930 MGTPELIDRECKMPGLSLGLADSOIKASFEWGHWOPLARLHNGSSTINAWIT-DKFS 1988
Oy 1209 -----WIKVDLAPMIHIGITOGAROKFSSLYISQITIMYSLDGKMQTYRGNSGT 1261
Db 1989 GSNKSPWIOVDMQREVFTGIOTGAKYLYKSYTTEFNVAVSSDQRMNRIEFGNSTKN 2048
Oy 1262 LMFVGNVSSSGKHNIFNPITARIIRLHPYISIRSTLRMELMGCDLMSGPIGME5 1321
Db 2049 VWFNGNSDASTTENQFOPPVARYIRISPTESYNPKARLELQCEVINGCSTPIGME5 2108
Oy 1322 KATSDAQITASSYFTNMFEAT-MSPSKARLHLOGRSNAMPQVNNPKMLVDFOKTKMT 1380
Db 2109 GNIKNGQITASSFKKSKMGWDYEPFRARLNAQGRVANMAKANNONMLOIDLKIKIT 2168
Oy 1381 GVTTOGVKSLTSMYKFEKLSSSDQGHOWTLFQNGV--KVFQGNDSFTTPVNSLDP 1438
Db 2169 ATTOGCKSLSEPMYKRYTIQYSDRGVEKSYREKSSMYDKIFEENNNIKGHYKNFNP 2228
Oy 1439 PLTTRRLRTHPOSWHQIALRMEVLGCEADLY 1471
Db 2229 PIIISREIRIIPKWNOSIALRLELFGC---DIY 2258

RESULT 6
O92024 PRELIMINARY; PRT: 1157 AA.
AC O92024.
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE HEPHAESTIN.
GN HEPH.
OC Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA MEDLINE=99140771; PubMed=9988272;
RA Vulpe C.D., Kuo Y.M., Murphy T.L., Cowley L., Askwith C., Libina N.,
RA Gitschier J., Anderson G.J.;
RT "Hephæstin, a ceruloplasmin homologue implicated in intestinal iron
RT transport, is defective in the sla mouse.";
RL Nat. Genet. 21:195-199(1999).
DR EMBL: AF082567; AAD16035.1; -.
DR HSSP: P00450; 1KCM.
DR MGD: MGI:1332340; Heph.
DR InterPro: IPR001117; -.
DR InterPro: IPR002355; -.
DR Pfam: PF00394; Cu-oxidase; 3.
DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 3.
DR PROSITE: PS00080; MULTICOPPER_OXIDASE; 1.
SO SEQUENCE 1157 AA; 129681 MW; 411BD1DA2BFEC2FD CRC64;
```

```
Query Match 22.8%; Score 1792; DB 11; Length 1157;
Best Local Similarity 33.8%; Pred. No. 7e-130;
Matches 402; Conservative 171; Mismatches 418; Indels 200; Gaps 20;

Oy 20 ATRHYIGAVELMSDYQOSDLGELPDARPPVPKSF-----PENTSVYKKTLEVEFV 75
Db 24 AIRNYYLIGIDMQMNAAPKGRNVITNOTLNDTVASSFLSKGRNIRGSSKYKTYKEYSD 83
Oy 76 HLENIARPPMMGLIGPTIOAEVYDPTVITLKNMASHPLAVGVWAKSEGEAYD 135
Db 84 GTYTEELAKPAWLGFLGPLLQAEVGYILHLKNFASREPTTHPHGVFEKSEGSLYD 143
Oy 136 QTSOREKEDDKVPPGGSHYVWQVLENGPMASDPLCLTYSYSHVDLVKNSLIGAL 195
Db 144 GSSGYLTKADDSVPPGSHVYVNSIPESHAFTPADPACTIWIYHSHVADAPDIATGICPL 203
Oy 196 LVNREGSL-----AKEKQTLHKFILLFAVDEKSMHSE-----TKNLMODRQAASR 245
Db 204 ITCKRGTLDSNPPQRKVDVHNEFLFVSIDENLSMHLNDNIATYCSDPASVDKEDGAQ 263
Oy 246 AMPKMTVNGVYNSRLPGLIGCHRSVYMHVIGMGTTPVHSHFEGHFLVRNHRASL 305
Db 264 DSNRMAHINGFVGNLPELSMCAQKHVAHNLFGMGNEIDVHTAFFGQMLSTRGHTTVA 323
Oy 306 EISPTFLTAQTLMDLGOFLFCHLSSHQHDGEAYVYKVSCEPBPOLRMKNNEABDY 365
Db 324 NIFPATFVTAEMVPQKSGTWLSCENVSHLRSQMAFYKYKVSMDPPV----- 372
Oy 366 DDOLITSEMVYAFDDNSPSFIOISVAKKHKTWVHVAIEEEDWMDAPLYLAPDRS 425
Db 373 -DOLTG-----KVROYLQAEHLOMDGP--IGYDGT 402
Oy 426 YKS-----OYLNNGPORIGKRYKVFMAVYDTEFTKTRALQHS--GILGPLYNG 474
Db 403 GKSLREBGSPPDKYFOKSSSRIGTYWAKYAEFOBETQOEYHQBEEHLGILGPVIRA 462
Oy 475 EVGDTLLIFKQNASRPNYIPIHGTIVRPLYSRLPKGVKH---LKDPILGGEIFYK 531
Db 463 EVGDTIQVVEYNRASQPFSLQPHGVFEKNESEGTAVNDGTSHPKVAKSP-----EKVY 517
Oy 532 WTVTVEDEGPKSPRCLTRYSFVMMERDLASGLGPLLICYKESVQORGNQMSDKRN 591
Db 518 WTVPPHAGPTADPACLTWYFSAADPTDNTGSLGPLLVCAKAGLGDGKQKGVDEF 577
Oy 592 VILFSEVDENRSMYITENTIOFLPNPAG-----VOLEDE-FOASNIMHSGVYFDS 644
Db 578 FLLEFYFDEHESWYNNAN-----QAAGMDSLSLSEDEVGODSRMAHINGLFESNLP 631
Oy 645 QLSVCLHEVAAYTILSIGAOTFLSVFSGYTFKHMVYEDTLTLPFESGETVFMSENP 704
Db 632 RLDMCKGDTVAWHLLLGITETDVHGVFEGNTVOLGMRKGAVMLFPHFTVTAIQPNP 691
Oy 705 GLWILCHNSDFRNRGMTLLKYSQCDKNTGDYEDSYSDISAVYLSKNAALERPSSFQN 764
Db 692 GIFEIYCQAGSHREEGMQALYVNSQ-----SSHOD----- 722
Oy 765 SRHPSTROKOFNATPVYLKRNHOREITRTLQSDOEEDIDYDTISYEMKKEDEIYDEDEN 824
Db 723 ----- 722
Oy 825 QSPRSFOKTRHYFIAVERLMDY-----GMSSSPHYLNRRAOSGSGVQPK 870
Db 723 -SPRQHYQASRYVYIWAEEIEMDYCDRSMELEMHNTSEKDSYHVLFSNKGGLGSKYK 781
Oy 871 KYVFOEFTDGSFTQPLRGEINHLGLLPYIRAEVEDINMTFRNQASRPSTFS-SLI 929
Db 782 KAVFREYTDGTFRIIPRPRSGPEEHLIGLPLRGEVGDILLTVFKNKASRPSTIAHGV 841
Oy 930 SYEDDROGAEPKRNKVNENETKYFMKVOHNAHPKDEFDKAAVAFSDVDLEKDVSG 989
Db 842 ESNITGEPQAAE-----FGEVLTITQWNIPEKSGEPPSDACVSNITYSANDPIKMG 894
Oy 990 LIGPLLVCHTNTLNPAHQVYVQEFALFTITDETKSMYFTENNER-NCRAPCNIOMED 1048
```

| Db     | 985  | LYGGLVLCIRNGLILEBNGSRNDMDREFAFLFLIPDPNSWTLKENIATATYGPRESSHVNLKD | 95d           |
|--------|--|---|---------------|
| Oy     | 1049   | PTFEKNTREFAHNGYIMDTLPGLVMAODQRIRWTLKSGNSENIHSIHFSCHVEFTVRKE     | 110           |
| Db     | 955  | ATPLESNMHAHINKLYANLEGLTVYOGSERVAYWMLAMGQDTHDHTVHFHAESEFLYQNGQ   | 100           |
| Oy     | 1109   | EYKMAVLNLYGVGFETVEMLPKSGIMRWVECLLGEHLHAGSSTLEFLYSNK             | 1159          |
| Db     | 1015   | STRADVVDLFPGETFEVEMVANSNGTWMHCHVTDWHAGMETIFLYLSHE               | 1065          |
| RESULT | 7  |   |               |
|        | 09JL97   | PRELIMINARY;  | PRT; 1084 AA. |
| AC     | 09JL97;  |   |               |
| DT     | 01-OCT-2000 (TREMBlrel, 15, Created)                                 |   |               |
| DT     | 01-OCT-2000 (TREMBlrel, 15, Last sequence update)                    |   |               |
| DT     | 01-MAR-2001 (TREMBlrel, 16, Last annotation update)                  |   |               |
| DE     | GPI-ANCHORED CERULOPLASMIN.  |   |               |
| OS     | Rattus norvegicus (Rat).   |   |               |
| OC     | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;    |   |               |
| OC     | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus |   |               |
| OX     | NCBI_Taxid=10116;  |   |               |
| RN     | [1]  |   |               |
| RP     | SEQUENCE FROM N.A.   |   |               |
| RC     | STRAIN=SPAGUE DAWLEY; TISSUE=BRAIN;                                  |   |               |
| RX     | MEDLINE=20127919; PubMed=10660599;                                   |   |               |
| RA     | Patel B.N., Dunn R.J., David S.;                                     |   |               |
| RT     | "Alternative RNA Splicing Generates a Glycosylphosphatidylinositol-  |   |               |
| RT     | anchored Form of Ceruloplasmin in Mammalian Brain."                  |   |               |
| RL     | J. Biol. Chem. 275:4305-4310(2000).                                  |   |               |
| DR     | EMBL; AF202115; AAF34175.1; -.                                       |   |               |
| DR     | InterPro; IPR001117; -.  |   |               |
| DR     | InterPro; IPR002355; -.  |   |               |
| DR     | Pfam; PF00394; Cu-oxidase; 3.  |   |               |
| DR     | PROSITE; PS00079; MULTICOPPER_OXIDASE2; 1.                           |   |               |
| DR     | PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.                           |   |               |
| QO     | SEQUENCE 1084 AA: 123748 MW; 9506584154257C55 CRC64;                 |   |               |

|                       |                   |  |             |              |
|-----------------------|-------------------|--|-------------|--------------|
| Query Match           | 21.6%             | Score 1703;  | DB 11;      | Length 1084; |
| Best Local Similarity | 32.2%             | Pred. No. 5e-123;  |             |              |
| Matches 388;          | Conservative 190; | Mismatches 427;  | Indels 200; | Gaps 23;     |
| QY                    | 5                 | LSTCFEFLDLRFCSATRRYLIGAVELSDYMQ--SDLGEL--PVDAFRPRVKESEFNTS         | 62          |              |
|                       |                   | : : : : :    : : : : :    : : : : :    : : : : :                   |             |              |
| Db                    | 6                 | LSALLFL--HSSLALPRKHHYYIGITEAVMDVASSEMEKELLISVDREQSNFLRLNDRICG      | 64          |              |
| QY                    | 63                | VYVKLTLEVEFTVHLFNIAKPRPPMGILGPTIOAEYDVTYITLKNMAHSHVSLHANGV         | 122         |              |
|                       |                   | : : : : :    : : : : :    : : : : :    : : : : :    : : : : :      |             |              |
| Db                    | 65                | RKYKALISEYTDGTFGTTIDKPAWLGLGLPVYKAEVDKSVAHKFNFSRPTTHAGV            | 124         |              |
| QY                    | 123               | SYMKASEGAEDDQTSQREKEDDKFPGSGSHYVQVLKENSPMAADPLCLTYSYLSHD           | 182         |              |
|                       |                   | :    :    :    : : : : :    : : : : :    : : : : :    : : : : :    |             |              |
| Db                    | 125               | TYTKANEBAITPRDNTDQRADDKLFEPQOQIYLVLRA--NEPSRGSDSNCAVRIHSHVD        | 183         |              |
| QY                    | 183               | LVKDINSGLIGALLVCREGSLAKEKTQTL--HKFTLLFAVFDKSGMHSFTKNSLM----        | 236         |              |
|                       |                   | :       :       :    :    : : : : :    : : : : :    : : : : :      |             |              |
| Db                    | 184               | APKDIASLIGLPIELCKKGSLLHKEEENIDQEFVLMFVSVDENLSMYLEDNLTKTFCSPE       | 243         |              |
| QY                    | 237               | -QDRPAASARMPKHNHTNGVNSRLPCLIGCHKRSYVHWVIGMOTTPRVHSIFLEGHTF         | 295         |              |
|                       |                   | :    : : : :    : :                      : : : : :       :       : |             |              |
| Db                    | 244               | KYDKCNEDFQSSNRMYSLNGTFFGLPGLSIMCAEDRVKMYLFGMENEVDYSELPHGQAL        | 303         |              |
| QY                    | 296               | LVRNHQASLEISPTTFPLTAQTLIMDLQGLFCHISSHOHDGMAYAVVDSCPREPQRL          | 355         |              |
|                       |                   | :    :    :    :    :    :    :    :    :    :    :    :    :    : |             |              |
| Db                    | 304               | TSKNYHTDITNLFPRTLIDVSVVAQNGVMMLSCGNLNLKAGLDAPFQVRRG-----           | 356         |              |
| QY                    | 356               | MKNNEAEADYDDDLTDSENDVVRFPDDNSPSFTQIRSVAKKIPKPTWVNIYIAAEEEDMDYA     | 415         |              |
|                       |                   | :    :    :    :    :    :    :    :    :    :    :    :    :    : |             |              |
| Db                    | 357               | --NKPSP--DDDDIDRRV-----RH-----YIAAEETIWDVA                         | 385         |              |

|  |      |   |      |
|--|------|---|------|
| OY   | 416  | P-----LVLAPDDBSXYSOVLNNGOPKIGSKYKVFAPMTDETF-----KTRAI         | 461  |
| OY   | 416  | P-----LVLAPDDBSXYSOVLNNGOPKIGSKYKVFAPMTDETF-----KTRAI         | 461  |
| Db   | 386  | PSGDTFTGGENLUSLGSDSVFEPQ-----GATRIGSGYAKKLVRREYTDSTFNKRQKQPD  | 441  |
| OY   | 462  | QHEGILIGPRLLYGCVGDTLLIFKKNQASRPYNIYPRGIDVR-----PLYSRRLPKGVNHL | 517  |
| Db   | 442  | EHNLGILIGPRLYMAEVBGIIINVTNNKNGQPRLSIQPRGAVFTFENEGTYG---PGRBSK | 498  |
| OY   | 518  | KDPLPLGELIFKRYKMTVLEDEGPTKSDPRCLATRYSSPVNMBERDLASGLIGPLLICYES | 577  |
| Db   | 499  | QASHVAKRETFTEYEWTPKCEMGPTVADPVCLSKMYSGVDLTIKDIFGTGLIGMKICKGS  | 558  |
| OY   | 578  | VDQNGNQMDSKRVNILEFSVFEDNBSMYLENIQRPLRPAGVQLEDEPQASNTMHSIN     | 637  |
| Db   | 559  | LADGRKQDVDKKEFYLFATVLEVENESILLDDNIRMTTAPENVVDKEDDEPQESNNKHSN  | 618  |
| OY   | 638  | GYVDSLSQ-LSVCLHEAVAYUYIISGAQDPLFSVFSGYTFKHKMYEDDTLLTFPFSGET   | 696  |
| Db   | 619  | GFMTGNLPGLMCLGSEITWYLFPSAGNEADYHGITYSGWTYLSKSKERDPTALPFRHKSLT | 678  |
| OY   | 697  | VFNSENPNGLMILGCHNSDNFRNRGMTALLKVSSCDKNTGDEYEDYSEDIISAVLISKNAI | 756  |
| Db   | 679  | LLMTPTGEGSDVDECLTTHDTYGMQKQKTYVNC-----KQGEEDVTIL-----         | 722  |
| OY   | 757  | EPRSFQNSRHPSTROKQFNATPPVLKRRHOREITRTTLOSDEEIDYDDTISVEKKEDF    | 816  |
| Db   | 723  | -----   | 722  |
| OY   | 817  | DIYEDENQSPRSQKTRHYFLAFAVERLMDGMSSPHVLRNRAQSSV-----            | 866  |
| Db   | 723  | -----QGERTYUATAAVEVENDYSPSRDMEMELHILQEBQVNSAFLDKKEF           | 767  |
| OY   | 867  | ---POFKVVFOEFTDSFTQPLRYRCELNEHNLGLLPYIRAEVEDNIMWTERNQASRYS    | 923  |
| Db   | 768  | FISGKYKVVYRETTDSTFEEDQYKRRAREEHNLGLRPLHADVQGVKVAFAFNMASTRYS   | 827  |
| OY   | 924  | FYSSLISTEEDQROGAEPKRNRYK---PNEGTYYKVOHNMAPTKDEDFCKAMAYFSDV    | 980  |
| Db   | 828  | IHA-----HGVTKSSSTVPTPLPBEVRYTIWQIPERSGAGTDESPCLPMAYYSTV       | 877  |
| OY   | 981  | DLEKDVHSGILGPLVC---HTWTLMPARGROYTVQEFRLFTTIDEXKSWTFENMERN     | 1037 |
| Db   | 878  | DRVADLVSGILGPLLYCRKSXYKVFNPK-----KKMEVSLFLVLEVDENESWLDNINIFY  | 932  |
| OY   | 1038 | CRAPCNLOMEDPTFKENYRFHAINGYIMDTPLGLVAADQRIRWYLLSGMSNENHSIH     | 1097 |
| Db   | 933  | SDHEKYNKQNEBERIESKNHAINGMKFNGLQGLTHYADVENVYVYAMAGNEIDLHTYHF   | 992  |
| OY   | 1098 | SGHVFYVRKKEEXKMALYNLPGVFETVEMLRSKACIMRWECLEHNLHAGNSTLELYS     | 1157 |
| Db   | 993  | HGHFQYKHNHGINHSSDVFELPCTYQTLMEFROTPTGWTLLHCHVMDHINAGWVTYTVLP  | 1052 |
| OY   | 1158 | NKCOF 1162  |      |
| Db   | 1053 | NQASS 1057  |      |
| RESULT 8   |      |   |      |
| O9XT27 PRELIMINARY; PRT: 1048 AA.                                    |      |   |      |
| AC O9XT27;   |      |   |      |
| DT 01-NOV-1999 (TEMBrel. 12, created)                                |      |   |      |
| DT 01-NOV-1999 (TEMBrel. 12, last sequence update)                   |      |   |      |
| DT 01-OCT-2000 (TEMBrel. 15, last annotation update)                 |      |   |      |
| DE CERUOPLASMIN.   |      |   |      |
| OS Ovis aries (Sheep).   |      |   |      |
| OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |      |   |      |
| OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea; |      |   |      |
| OC Bovidae; Caprinae; Ovis.  |      |   |      |
| OX NCBI_TaxID=9940.  |      |   |      |
| FN [1]   |      |   |      |
| RP SEQUENCE FROM N.A.  |      |   |      |
| OC TISSUE=LIVER;   |      |   |      |



|        |             |  |                                     |         |
|--------|-------------|--|-------------------------------------|---------|
| QY     | 370         | TDSEMDVVRFPDDNSPFIQIRSAKKHNRKTHWYIAAEEDMDYAPL                | -----VLA                            | 420     |
| Db     | 109         | TG-----  | -----KVQRFETAHNFIQMDYGMHDSGTGNKLE   | 141     |
| QY     | 421         | PDDRSKSQYLNNRPORIGRKVKKVRPMATYDTEFKTREAIOHES--GILGPLYGEVD    | 478                                 |         |
| Db     | 142         | PG--SISQFKQSSSSNIGGYMKVREARODEFQFKMLLEDNRHIGLIGYIRAEVD       | 199                                 |         |
| QY     | 479         | TLLIFKNOASRPYNIYPHGITDVRPLYSRRLPKGVKNHLDPRILPGEIFK           | 532                                 |         |
| Db     | 200         | TIQVVFYKRASQPSMOPHGV----                                     | FEYKDYEGTV--YUNDGSSYPGLYAKPREKTYRRM | 252     |
| QY     | 533         | TVTVJEDGTKDRPCILTRYVSSFYNNMRDLASGLIGLILCYKESVDQRONOJMSKRV    | 592                                 |         |
| Db     | 253         | TVPRPHAGTADPRCLATYMMFESADPLRIDNNSGLYGLVLAICRAGAAGADOKGVKVEEF | 312                                 |         |
| QY     | 593         | ILFSEVDENRSMYUTENIQEFLRPAQVOT-----EDPE--FOASNMHSINKYVDSL-    | 644                                 |         |
| Db     | 313         | LLFVLIDENKSMYEN-----ANQAAAMLDRLSDIEDGFQDSNRMHAINGFLESNLP     | 365                                 |         |
| QY     | 645         | QLSVCLHEVAWYTLISGAQOTDLSVFFSGYTEKHKMYEDTLTLRPPSGETVEMSMENP   | 704                                 |         |
| Db     | 366         | RIDMCKDDTYAMNHLIGTETEDVAGVFOGNTVOLQCMRGKGAAMLRPHGFVMAIMQDNL  | 425                                 |         |
| QY     | 705         | GLMILIGHNSDFRNRGMTALLKYSOSCKNTGDYEDYSIEDISAYLLSKNNAIEPRFSQ   | 764                                 |         |
| Db     | 426         | GTEFELYQAGSHRAGRAKRAYIYNSQC-----                             |                                     | 451     |
| QY     | 765         | SRHSSTROKQFNAPRLPKHOREIRRTTQILOSQOEIYDDTISVEMKKEDFDIYEDEN    | 824                                 |         |
| Db     | 452         | -----PQHQ-----   |                                     | 455     |
| QY     | 825         | QSPRSFOKTRHFIATAVERLDYGMSS--PHVLRNPAQSGSY-----POFK           | 870                                 |         |
| Db     | 456         | ATPRQRYOARITYIIMAEVEYMDJCPRDSMERENHNOSEKDSYGLFLSNKODLGSRYK   | 515                                 |         |
| QY     | 871         | KVVRQEFIDSGFTQRLRGELNENIGLIGYIAEVEDNIMYTFRRQASPRYSFS--SLI    | 929                                 |         |
| Db     | 516         | KAVREYTDGTFRLPRPRGTREEHGILGPLIGVEYDILTYVEKKNASPRYSVNAHGVY    | 575                                 |         |
| QY     | 930         | SYEEDOROGAERPKNFKPRMETKYEFKQYONHNAPTKDEFDCAKAWYFSDVLEKDVHSG  | 989                                 |         |
| Db     | 576         | ESTYVWPLAE-----PEBYVTYQONINIERSGRPGNDACSVIYYSAVDPKIMYSG      | 628                                 |         |
| QY     | 990         | LIGPLVCHTNTLPRAHGROYVOERALFTTFDETKSWYFTENM--ERNCRAPCNTOMED   | 1044                                |         |
| Db     | 629         | LVGRLAICOKGILIEPHGGRSDMREFFALLRLIDENKSWLIEENVATHGSDODGSINTLD | 688                                 |         |
| QY     | 1049        | PTFEENYRPHANINGYIMDTLPGYLMADODKIRYUILLMSGSENNHISHPSGHVFYTRKE | 1100                                |         |
| Db     | 689         | ETFLSKMKHAINKLANRGLIMTQGEVAMTYMLMAGDVVDLHTHIEAESFLYRGE       | 748                                 |         |
| QY     | 1109        | EYKMAIYLYPGVEYEVEMLRPSKAGIYRVECLIGENHLAQMSTLFLVSNKQOTPLGMA   | 1166                                |         |
| Db     | 749         | NYRADVYVDLPRGTFPEVEYEVASNPGETWLMHCHTVDHNAAGMETLFLPVEFS       | 799                                 |         |
| QY     | 1169        | GHIDPQITLASGYQGMARPLARLHNSGSINAMSTKEPYSWIKYVDLLAPRIHIGITQGA  | 1224                                |         |
| Db     | 800         | EHLSPRLVITIKETEKAVPRDIE-----EGNMKMLGMQIPK--NVEMIASVLY-----   | 846                                 |         |
| QY     | 1229        | RQKFSLYISOFITMYSLDCKKQMYT                                    | 1255                                |         |
| Db     | 847         | -----AISVTLILVYLAJGQVWYQHR                                   | 868                                 |         |
| RESULT | 10          |  |                                     |         |
| ID     | 075659      |  |                                     |         |
| AC     | 075659      | PRELIMINARY;   | PRT:                                | 782 AA. |
| DT     | 01-NOV-1998 | (TREMblrel. 08, Created)                                     |                                     |         |
| DT     | 01-NOV-1998 | (TREMblrel. 08, last sequence update)                        |                                     |         |
| DT     | 01-MAV-2000 | (TREMblrel. 13, last annotation update)                      |                                     |         |

[illegible]

```
QY 754 NAIEPRFSQNSRHPSTROKQENAPPVLKRHOREITRTTLOSQDEIDYDPTISVEMKK 813
D 647 -----PGHQ-----
QY 814 EDFDIYDEDENSPSPFOKTRHYFIAVERBLMDYGMSSS-PAVLNKRAGSGV----- 866
D 651 -----ATPRROGAARIYYIMAEVEEDYCPDRSWEREMWNSKDSYGYFLS 699
QY 867 -----POFKVVFQEFIDGFTOPLYRGELNHLGLGPYIRAEVDNIWTFRRQAS 919
D 700 NKDGLGSRKKAAVREYIDGTFRRIPRPTGPEHLGILIKGEVGDILYVFKNAS 759
QY 920 RPYSFYS 926
D 760 RPYSVHA 766

RESULT 11
Q14286 PRELIMINARY; PRT; 216 AA.
ID Q14286
AC Q14286;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE COAGULATION FACTOR VIII ASSOCIATED PROTEIN B.
GN F8B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93052386; PubMed=1427887;
RA Levinson B., Kendrick S., Gamel P., Fisher K., Gitschier J.;
RT "Evidence for a third transcript from the human factor VIII gene.";
RL Genomics 14:585-589(1992).
DR EMBL; M90707; AAC58466.1; -.
DR HSPB; P00451; ICFG.
DR InterPro; IPR000421; -.
DR Pfam; PF00754; F5_F8_type_C; 1.
DR PROSITE; PS01285; FA58C_1; UNKNOWN_1.
DR PROSITE; PS01286; FA58C_2; 2.
DR SMART; SM00231; FA58C; 1.
SQ SEQUENCE 216 AA; 24641 MW; 6C82D4F89E35A376 CRC64;

Query Match 14.0%; Score 1104; DB 4; Length 216;
Best Local Similarity 100.0%; Pred. No. 1,le-77;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1264 VFFGVNDSSGIKHNIFNPPIARIYRIHLPTHTYSIRSTLRMELMGCDLNSCSPMLGMSKA 1323
D 9 VFFGVNDSSGIKHNIFNPPIARIYRIHLPTHTYSIRSTLRMELMGCDLNSCSPMLGMSKA 68
QY 1324 ISDAQITASSYFTNMFAWSPSKARLHLGSRNAMPQVNNPKEMQLQVDFOKTKVTGYT 1383
D 69 ISDAQITASSYFTNMFAWSPSKARLHLGSRNAMPQVNNPKEMQLQVDFOKTKVTGYT 128
QY 1384 TQGVKSLTSMYVKEFLISSSDGHQWTLFFQNGKVKYVQGNDSFTPVVNSLDPPLLTR 1443
D 129 TQGVKSLTSMYVKEFLISSSDGHQWTLFFQNGKVKYVQGNDSFTPVVNSLDPPLLTR 188
QY 1444 YLRHPQSWHQAIALRMEVLGCEADLY 1471
D 189 YLRHPQSWHQAIALRMEVLGCEADLY 216

RESULT 12
Q9RLX9 PRELIMINARY; PRT; 463 AA.
AC Q9RLX9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
```

```
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE MILK FAT GLOBULE GLYCOPROTEIN MFG-E8 LONG FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE=MAMMARY GLAND;
RX MEDLINE=99120894; PubMed=9920772;
RA Oshima K., Aoki N., Negi M., Kishi M., Kitajima K., Matsuda T.;
RT "Lactation-dependent expression of an mRNA splice variant with an exon
RT for a multiply O-glycosylated domain of mouse milk fat globule
RT glycoprotein MFG-E8."
RL Biochem. Biophys. Res. Commun. 254:522-528(1999).
DR EMBL; AB021130; BAA35180.1; -.
DR HSPB; P00740; IIXA.
DR InterPro; IPR000421; -.
DR InterPro; IPR000561; -.
DR InterPro; IPR001092; -.
DR InterPro; IPR001438; -.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR PRINTS; PR00010; EGFBL00D.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR SMART; SM00181; EGF; 1.
KW EGF-like domain; Glycoprotein.
SQ SEQUENCE 463 AA; 51269 MW; D719D2BE090E6427 CRC64;

Query Match 8.4%; Score 664; DB 11; Length 463;
Best Local Similarity 42.5%; Pred. No. 5,le-43;
Matches 136; Conservative 62; Mismatches 110; Indels 12; Gaps 5;

QY 1157 SNKCOTPLMASGHIRDFOITASGOY-----GONAPKLARHYSGSINAW--STKEPFS 1208
D 145 ASKCSITQLMGEGALINDSISASVYMGLOIRNGCEPLARIYRGIYANAWASVYDSKP 204
QY 1209 WIKVLDLAPMIHIGIKTGAROKFSSLYISOPTIMYSLDGKKWQTYRGNSGTGLWVFFGN 1268
D 205 WIQVNLRLKRMVSGVMTQASRASGRAEYLTKEKVAVYSLDGKRFELIDESGD-KEFLGN 263
QY 1269 VDSGSGIKHNIFNPPIARIYRIHLPTHTYSIRSTLRMELMGCDLNSCSPMLGMSKAISDAQ 1328
D 264 LDNNSLTKVMEFNPTEAQTIRLYPVSCHRGCTLRRELLGCEHLGSGSEPLGKNNTPDSQ 323
QY 1329 ITASSYFT--NMFA-TWSPSKARLHLGSRNAMPQVNNPKEMQLQVDFOKTKVTGYTQ 1385
D 324 MSASSSYKTWNLRAPFGVYHLGRDNGKINAWTQNSAKKMLQVLDGTQRQYGIITL 383
QY 1386 GYKSLTSMYVKEFLISSSDGHQWTLFFQNGKVKYVQGNDSFTPVVNSLDPPLLTRYL 1445
D 384 GARDFGHIQYVASYKVAHSDDGQWTVYEEOGSSKVFQGLDNNSHKKNIFEKPMARYV 443
QY 1446 RHHPQSWHQAIALRMEVLG 1465
D 444 RVLPSVMHNRITRLLELGC 463

RESULT 13
Q43854 PRELIMINARY; PRT; 480 AA.
AC Q43854;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE INTEGRIN-BINDING PROTEIN DEL1 PRECURSOR.
GN DEL1.
OS Homo sapiens (Human).
```





```
DB 206 DGRKEFFIDODESGD-KEFLGNLDNNSLRVNMENPTLEAOYLRLPYVSCRGCTLRPELL 264
OY 1307 GCDLNSCMLPLMESKAISDAOTASSYFT--NMFA-TNSPSPARHLLOGRSNAMPPOVN 1363
DB 265 GCELHGCSPEPLGLKNNITIDSOHSASSSKYTNLRAFQWPHLGRDLNCKITNAWTAQSN 324
OY 1364 NPKEMLOVDFOFKMYKVTGTVGKSLTSMVYKEFLISSODGHQWTLFFONGKVKVFO 1423
DB 325 SAKEMLOVDLGTQROVYTGITGARDFGHILQYVASYKVAHSDQGVQWTVYEEOGSSKVFQ 384
OY 1424 GNQDSFTPVNSLDPLLTLYRLIHPQSWHQIALRMEVIGC 1465
DB 385 GNLDNNSHKNIFEKPFMARVYRVLPVSWHNRITLRLLELGC 426

RESULT 15
ID 035474 PRELIMINARY; PRT: 480 AA.
AC 035474;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, last annotation update)
DE INTEGRIN-BINDING PROTEIN DELT PRECURSOR.
GN EDL3 OR DEL1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND CHARACTERIZATION.
RC TISSUE=EMBRYO.
RA MEDLINE=98083109; PubMed=9420328;
RA Hital C., Zupancic T., Penta K., Mikhail A., Kawana M.,
RA Quettermous E.E., Aoka Y., Fukagawa M., Matsui Y., Platika D.,
RA Auerbach R., Hogan B.L.M., Snodgrass R., Quettermous T.;
RT "Cloning and characterization of developmental endothelial locus-1: an
RT embryonic endothelial cell protein that binds the alphabeta3 integrin
RL receptor.";
RL Genes Dev. 12:21-33(1998).
CC -1- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH
CC INTERACTION WITH THE ALPHA-V BETA INTEGRIN RECEPTOR. INHIBITS
CC FORMATION OF VASCULAR-LIKE STRUCTURES. MAY BE INVOLVED IN
CC REGULATION OF VASCULAR MORPHOGENESIS OR REMODELING IN EMBRYONIC
CC DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.
CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS: LONG ISOFORM (SHOWN HERE) AND
CC SHORT ISOFORM, ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ANGIOBLASTS AND EARLY ENDOTHELIAL
CC CELLS. BY EMBRYONIC DAY 13.5, ALSO EXPRESSED IN A RESTRICTED GROUP
CC OF NON-ENDOTHELIAL CELLS INCLUDING CHONDROCYTES AND RETINAL
CC NEURONS.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE EMBRYO FROM DAY 7. AFTER DAY
CC 15.5, EXPRESSION DECREASES AND DISAPPEARS COMPLETELY BY THE TIME
CC OF BIRTH.
CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC EMBL: AF031524; AAB86585.1; -.
DR EMBL: AF031525; AAB86586.1; -.
DR HSSP: P00740; 1EDM.
DR MGD: MGI:1329025; Edl13.
DR InterPro: IPR000152; -.
DR InterPro: IPR000421; -.
DR InterPro: IPR000561; -.
DR InterPro: IPR000742; -.
DR InterPro: IPR001881; -.
DR Pfam: PF00008; EGF_3.
DR Pfam: PF00754; F5_F8_type_C_2.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_3.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01187; EGF_CA; 1.
DR PROSITE: PS01285; FA58C_1; 2.
DR PROSITE: PS01285; FA58C_1; 2.
```

```
DR PROSITE: PS01286; FA58C_2; 2.
DR SMART: SM00231; FA58C; 1.
KW EGF-like domain; Alternative splicing; signal; Developmental protein;
KW Cell adhesion; Repeat; Vascularization.
FT SIGNAL 1 16
FT CHAIN 17 480
FT DOMAIN 26 59
FT DOMAIN 78 116
FT DOMAIN 123 154
FT DOMAIN 161 311
FT DOMAIN 322 473
FT SITE 96 98
FT DISULFID 26 37
FT DISULFID 31 48
FT DISULFID 50 59
FT DISULFID 78 89
FT DISULFID 83 105
FT DISULFID 107 116
FT DISULFID 158 314
FT DISULFID 301 305
FT DISULFID 319 476
FT VARSPIC 218 221
FT VARSPIC 222 480
SQ SEQUENCE 480 AA; 53740 MW; 4CD91EE9261714D CRC64;
```

Query Match 8.3%; Score 655.5; DB 11; Length 480;  
Best local Similarity 44.6%; Pred. No. 2.5e-42;  
Matches 145; Conservative 45; Mismatches 120; Indels 15; Gaps 5;

```
OY 1159 KCOTPLGMSGHRIDPOITRASGOY-----GQAPPLALRLHSGSINAMSTE--PFSWI 1210
DB 157 KCSGPIGIEGIIISNOOITASSSTHRAFLGLRKYPYARLNKGLINAWTAENDRWPI 216
OY 1211 KVDLAPMITHIGIKTGAROKFSLYISOFILYSLDGKKMOTYRGNSTGLMFEFGVND 1270
DB 217 QINDLRKMRVYTGITGAKRIGSPETIKSYKIAYSNDGKTWAMYKKKGTNEWVRGNDV 276
OY 1271 SSGIKHNIFNPPIIARYIRLPHYSIRSTLRLMELGCDLNSCMLPLMESKAISDAOT 1330
DB 277 NNPVANSFPPIKQYVRLPQICRRHCTLRRELGLCGLSCSPGLMKSGHIDYQIT 336
OY 1331 ASSYF---TNMFATSPSKARLHLOGRSNAMPPOVNNKEWLOVDFOKTKMYTGVTTQG 1386
DB 337 ASSVFTLNNDMF-TWEPKARLDKQKVNAMTSGHDSQWLOVDLVPVTGTIITQG 395
OY 1387 VKSLTSMVYKEFLISSODGHQWTLF--FONGKVVFPQGNDSFTPVVNSLDPLTRY 1444
DB 396 AKDFGHVQFVSGSKLYISNDGEHMYHODEKQKDKVFGQNFNDNTHRKNVYIDPPIYARF 455
OY 1445 LRTHPOSWVHQIALRMEVLGCEAOD 1469
DB 456 IRLPMSWYGRITLRSELGCAEE 480
```

Search completed: November 17, 2001, 13:25:24  
Job time: 188 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2001, 13:22:16 ; Search time 16.43 seconds  
(without alignments)  
3066.940 Million cell updates/sec

Title: US-09-689-430-2  
Perfect score: 7869  
Sequence: 1 MJELSTCFELCLRFCSA.....WVHQLRMELVGEADLY 1471

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues  
Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | ID | Description        |
|------------|--------|-------------|--------|----|--------------------|
| 1          | 7412   | 94.2        | 2351   | 1  | FA8_HUMAN          |
| 2          | 6347   | 80.7        | 2133   | 1  | FA8_PIG            |
| 3          | 6301   | -80.1       | 2319   | 1  | FA8_MOUSE          |
| 4          | 2405.5 | 30.6        | 2224   | 1  | FA5_HUMAN          |
| 5          | 2379   | 30.2        | 2211   | 1  | FA5_BOVIN          |
| 6          | 1725   | 21.9        | 1065   | 1  | CERU_HUMAN         |
| 7          | 1678   | 21.3        | 1059   | 1  | CERU_RAT           |
| 8          | 1639   | 20.8        | 1062   | 1  | CERU_MOUSE         |
| 9          | 663    | 8.4         | 427    | 1  | MFGM_RAT           |
| 10         | 657    | 8.3         | 463    | 1  | MFGM_MOUSE         |
| 11         | 650    | 8.3         | 409    | 1  | MFGM_PIG           |
| 12         | 635    | 8.1         | 427    | 1  | MFGM_BOVIN         |
| 13         | 588    | 7.5         | 387    | 1  | MFGM_HUMAN         |
| 14         | 469.5  | 6.0         | 931    | 1  | MRP2_HUMAN         |
| 15         | 464.5  | 5.9         | 925    | 1  | MRP2_RAT           |
| 16         | 462.5  | 5.9         | 931    | 1  | MRP2_MOUSE         |
| 17         | 458.5  | 5.8         | 914    | 1  | MRP1_CHICK         |
| 18         | 451.5  | 5.7         | 922    | 1  | MRP1_RAT           |
| 19         | 446.5  | 5.7         | 923    | 1  | MRP1_MOUSE         |
| 20         | 443    | 5.6         | 928    | 1  | MRP1_XENLA         |
| 21         | 429.5  | 5.5         | 923    | 1  | MRP1_HUMAN         |
| 22         | 306.5  | 3.3         | 3133   | 1  | HMCT_BOVMO         |
| 23         | 260.5  | 3.3         | 280    | 1  | XRRL_FUGRU         |
| 24         | 247    | 3.1         | 224    | 1  | XRRL_MOUSE         |
| 25         | 243    | 3.1         | 224    | 1  | XRRL_HUMAN         |
| 26         | 198.5  | 2.5         | 854    | 1  | DDR2_MOUSE         |
| 27         | 192    | 2.4         | 913    | 1  | DDR1_HUMAN         |
| 28         | 185.5  | 2.4         | 855    | 1  | DDR2_HUMAN         |
| 29         | 185.5  | 2.4         | 911    | 1  | DDR1_RAT           |
| 30         | 149.5  | 1.9         | 911    | 1  | DDR1_MOUSE         |
| 31         | 142    | 1.8         | 578    | 1  | ASO_TOBAC          |
| 32         | 135.5  | 1.7         | 1196   | 1  | BXCN_SCHPO         |
| 33         |        |             |        |    | P46081 clostridium |

## ALIGNMENTS

| RESULT | ID   | FA8_HUMAN                         | STANDARD: | PRT: | 2351 AA. |
|--------|--|-----------------------------------|-----------|------|----------|
| AC     | P00451   |                                   |           |      |          |
| DT     | 21-JUL-1986  | (Rel. 01, Created)                |           |      |          |
| DT     | 21-JUL-1986  | (Rel. 01, Last sequence update)   |           |      |          |
| DT     | 01-OCT-2000  | (Rel. 40, Last annotation update) |           |      |          |
| DE     | COAGULATION FACTOR VIII PRECURSOR (PROCOAGULANT COMPONENT)             |                                   |           |      |          |
| DE     | (ANTHEMOPHILIC FACTOR) (AHF).  |                                   |           |      |          |
| GN     | F8 OR F8C.   |                                   |           |      |          |
| OS     | Human sapiens (Human).   |                                   |           |      |          |
| OC     | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;      |                                   |           |      |          |
| OC     | Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.              |                                   |           |      |          |
| OX     | NCBI_TaxID=9606;   |                                   |           |      |          |
| RN     | [1]  |                                   |           |      |          |
| RP     | SEQUENCE FROM N.A.   |                                   |           |      |          |
| RX     | MEDLINE=86081164; PubMed=3935400;                                      |                                   |           |      |          |
| RA     | Treut M.A., Blacher R., Burke R.L., Caput D., Chu C., Dina D.,         |                                   |           |      |          |
| RA     | Harlog K., Kuo C.H., Maslarz F.R., Merryweather J.P., Najarian A.,     |                                   |           |      |          |
| RA     | Pachl C., Potter S.J., Puma J., Quiroga M., Rall L.B., Randolph A.,    |                                   |           |      |          |
| RA     | Urdea M.S., Valenzuela P., Dahl H.-H.M., Favalaro J., Hansen J.,       |                                   |           |      |          |
| RA     | Norfang O., Ezban M.,  |                                   |           |      |          |
| RT     | "Characterization of the polypeptide composition of human factor       |                                   |           |      |          |
| RT     | VIII:C and the nucleotide sequence and expression of the human kidney  |                                   |           |      |          |
| RT     | cDNA.";  |                                   |           |      |          |
| RL     | DNA 4:333-349(1985).   |                                   |           |      |          |
| RN     | [2]  |                                   |           |      |          |
| RP     | SEQUENCE FROM N.A.   |                                   |           |      |          |
| RX     | MEDLINE=85061548; PubMed=6438526;                                      |                                   |           |      |          |
| RA     | Wood W.I., Capon D.J., Simonsen C.C., Eaton D.L., Gitschler J.,        |                                   |           |      |          |
| RA     | Key B., Seeburg P.H., Smith D.H., Hollingshead P., Wilson K.L.,        |                                   |           |      |          |
| RA     | Delwart E., Tuddenham E.G.D., Vehar G.A., Lawn R.M.;                   |                                   |           |      |          |
| RT     | "Expression of active human factor VIII from recombinant DNA clones."; |                                   |           |      |          |
| RL     | Nature 312:330-337(1984).  |                                   |           |      |          |
| RN     | [3]  |                                   |           |      |          |
| RP     | SEQUENCE FROM N.A.   |                                   |           |      |          |
| RX     | MEDLINE=85061550; PubMed=6438528;                                      |                                   |           |      |          |
| RA     | Toole J.J., Knopf J.L., Wozney J.M., Sultzman L.A., Buecker J.L.,      |                                   |           |      |          |
| RA     | Pittman D.D., Kaufman R.J., Brown E., Shoemaker C., Orr E.C.,          |                                   |           |      |          |
| RA     | Amphlett G.W., Foster W.B., Coe M.L., Knutson G.J., Pass D.N.,         |                                   |           |      |          |
| RT     | Hewick R.M.;   |                                   |           |      |          |
| RT     | "Molecular cloning of a cDNA encoding human antihemophilic factor.";   |                                   |           |      |          |
| RL     | Nature 312:342-347(1984).  |                                   |           |      |          |
| RN     | [4]  |                                   |           |      |          |
| RP     | SEQUENCE FROM N.A.   |                                   |           |      |          |
| RX     | MEDLINE=93265012; PubMed=1303178;                                      |                                   |           |      |          |
| RA     | Gitschler J., Wood W.I.;   |                                   |           |      |          |
| RT     | "Sequence of the exon-containing regions of the human factor VIII      |                                   |           |      |          |
| RT     | gene.";  |                                   |           |      |          |
| RL     | Hum. Mol. Genet. 1:199-200(1992).                                      |                                   |           |      |          |
| RN     | [5]  |                                   |           |      |          |
| RP     | SEQUENCE OF 2064-2070 FROM N.A.  |                                   |           |      |          |
| RA     | de Water N.S., Williams R., Browett P.J.;                              |                                   |           |      |          |
| RL     | Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.                |                                   |           |      |          |
| RN     | [6]  |                                   |           |      |          |
| RP     | SULFATATION OF 1699.   |                                   |           |      |          |

RX MEDLINE-91093266; PubMed-1898735;  
RA Leyte A., van Schijndel H.B., Niehrs C., Huttner W.B., Verbeet M.P.,  
RA Mertens K., van Mourik J.A.;  
RT "Sulfation of Tyr1680 of human blood coagulation factor VIII is  
RT essential for the interaction of factor VIII with von Willebrand  
RT factor";  
RL J. Biol. Chem. 266:740-746(1991).  
RN [17]  
RP SULFATATION.  
RX MEDLINE-92207952; PubMed-1554716;  
RA Pittman D.D., Wang J.H., Kaufman R.J.;  
RT "Identification and functional importance of tyrosine sulfate  
RT residues within recombinant factor VIII.";  
RL Biochemistry 31:3315-3325(1992).  
RN [18]  
RP STRUCTURE BY NMR OF 2322-2343.  
RX MEDLINE-95200924; PubMed-7893714;  
RA Gilbert G.E., Baleja J.D.;  
RT "Membrane-binding peptide from the C2 domain of factor VIII forms an  
RT amphipathic structure as determined by NMR spectroscopy.";  
RL Biochemistry 34:3022-3031(1995).  
RN [19]  
RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.  
RX MEDLINE-91221499; PubMed-1902642;  
RA Gitschler J.;  
RT "The molecular basis of hemophilia A.";  
RL Ann. N.Y. Acad. Sci. 614:89-96(1991).  
RN [10]  
RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.  
RX MEDLINE-89088506; PubMed-2491949;  
RA White G.C. II, Shoemaker C.B.;  
RT "Factor VIII gene and hemophilia A.";  
RL Blood 73:1-12(1989).  
RN [11]  
RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.  
RX MEDLINE-95245332; PubMed-7728145;  
RA Antonarakis S.E., Kazazian H.H., Tuddenham E.G.D.;  
RT "Molecular etiology of factor VIII deficiency in hemophilia A.";  
RL Hum. Mutat. 5:1-22(1995).  
RN [12]  
RP VARIANT GLN-2326.  
RX MEDLINE-86235434; PubMed-3012775;  
RA Gitschler J., Wood W.I., Shuman M.A., Lawn R.M.;  
RT "Identification of a missense mutation in the factor VIII gene of a  
RT mild hemophilic.";  
RL Science 232:1415-1416(1986).  
RN [13]  
RP VARIANT PRO-2135.  
RX MEDLINE-88096539; PubMed-3122181;  
RA Levinson B., Janco R.L., Phillips J.A. III, Gitschler J.;  
RT "A novel missense mutation in the factor VIII gene identified by  
RT analysis of amplified hemophilia DNA sequences.";  
RL Nucleic Acids Res. 15:9797-9805(1987).  
RN [14]  
RP VARIANT GLN-2228.  
RX MEDLINE-88191889; PubMed-2833855;  
RA Yousoufian H., Antonarakis S.E., Bell W., Griffin A.M.,  
RT "Nonsense and missense mutations in hemophilia A: estimate of the  
RT relative mutation rate at CG dinucleotides.";  
RL Am. J. Hum. Genet. 42:718-725(1988).  
RN [15]  
RP VARIANT GLY-291.  
RX MEDLINE-88220354; PubMed-2835904;  
RA Yousoufian H., Wong C., Aronis S., Platokoukis H., Kazazian H.H. Jr.,  
RA Antonarakis S.E.;  
RT "Moderately severe hemophilia A resulting from Glu->Gly substitution  
RT in exon 7 of the factor VIII gene.";  
RL Am. J. Hum. Genet. 42:867-871(1988).  
RN [16]  
RP VARIANT CYS-1708.  
RX MEDLINE-89274393; PubMed-2499363;  
RA O'Brien D.P., Tuddenham E.G.;

RT "Purification and characterization of factor VIII 1,689-Cys: a  
RT nonfunctional cofactor occurring in a patient with severe hemophilia  
RT A.";  
RL Blood 73:2117-2122(1989).  
RN [17]  
RP VARIANT CYS-391.  
RX MEDLINE-90001543; PubMed-2506948;  
RA Shima M., Ware J., Yoshioka A., Fukui H., Fulcher C.A.;  
RT "An arginine to cysteine amino acid substitution at a critical  
RT thrombin cleavage site in a dysfunctional factor VIII molecule.";  
RL Blood 74:1612-1617(1989).  
RN [18]  
RP VARIANT LEU-189.  
RX MEDLINE-90057680; PubMed-2510835;  
RA Chan V., Chan T.K., Tong T.M., Todd D.;  
RT "A novel missense mutation in exon 4 of the factor VIII:c gene  
RT resulting in moderately severe hemophilia A.";  
RL Blood 74:2688-2691(1989).  
RN [19]  
RP VARIANT LEU-2326.  
RX MEDLINE-89197216; PubMed-2495245;  
RA Inaba H., Fujimaki M., Kazazian H.H. Jr., Antonarakis S.E.;  
RT "Mild hemophilia A resulting from Arg-to-Leu substitution in exon 26  
RT of the factor VIII gene.";  
RL Hum. Genet. 81:335-338(1989).  
RN [20]  
RP VARIANT HIS-391.  
RX MEDLINE-89264602; PubMed-2498882;  
RA Arai M., Inaba H., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,  
RA Fujimaki M., Hoyer L.W.;  
RT "Direct characterization of factor VIII in plasma: detection of a  
RT mutation altering a thrombin cleavage site  
RT (arginine-372->histidine).";  
RL Proc. Natl. Acad. Sci. U.S.A. 86:4277-4281(1989).  
RN [21]  
RP VARIANT CYS-1708.  
RX MEDLINE-90105723; PubMed-2104766;  
RA Arai M., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,  
RA Phillips J.A. III, Janco R.L., Hoyer L.W.;  
RT "Characterization of a thrombin cleavage site mutation (Arg 1689 to  
RT Cys) in the factor VIII gene of two unrelated patients with  
RT cross-reacting material-positive hemophilia A.";  
RL Blood 75:384-389(1990).  
RN [22]  
RP VARIANTS GLN-2228 AND LEU-2326.  
RX MEDLINE-90123183; PubMed-2105106;  
RA Casula L., Murru S., Pecorara M., Ristaldi M.S., Restagno G.,  
RA Mancuso G., Morfini M., de Biasi R., Baudo F., Carbonara A.;  
RT "Recurrent mutations and three novel rearrangements in the factor  
RT VIII gene of hemophilia A patients of Italian descent.";  
RL Blood 75:662-670(1990).  
RN [23]  
RP VARIANT CYS-391.  
RX MEDLINE-90329422; PubMed-1973901;  
RA Patinson J.K., McVey J.H., Boon M., Ajani A., Tuddenham E.G.;  
RT "CRM+ hemophilia A due to a missense mutation (372->Cys) at the  
RT internal heavy chain thrombin cleavage site.";  
RL Br. J. Haematol. 75:73-77(1990).  
RN [24]  
RP VARIANTS PHE-1699 AND CYS-1708.  
RX MEDLINE-90152691; PubMed-2105906;  
RA Higuchi M., Wong C., Kochan L., Olek K., Aronis S., Kasper C.K.,  
RA Kazazian H.H., Antonarakis S.E.;  
RT "Characterization of mutations in the factor VIII gene by direct  
RT sequencing of amplified genomic DNA.";  
RL Genomics 6:65-71(1990).  
RN [25]  
RP VARIANTS CYS-1728 AND ASP-1941.  
RX MEDLINE-90169988; PubMed-2106480;  
RA Traystman M.D., Higuchi M., Kasper C.K., Antonarakis S.E.,  
RA Kazazian H.H.;  
RT "Use of denaturing gradient gel electrophoresis to detect point  
RT mutations in the factor VIII gene.";









OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;  
 RX MEDLINE=93300511; PubMed=8314577;  
 RA Elder B., Lachich D., Gletscher J.;  
 RT "Sequence of the murine factor VIII cDNA";  
 RL Genomics 16:374-379(1993).  
 CC -1- FUNCTION: FACTOR VIII, ALONG WITH CALCIUM AND PHOSPHOLIPID, ACTS  
 CC AS A COFACTOR FOR FACTOR IXA WHEN IT CONVERTS FACTOR X TO THE  
 CC ACTIVATED FORM, FACTOR XA.  
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.  
 CC -1- TISSUE SPECIFICITY: FOUND IN MOST TISSUES.  
 CC -1- SIMILARITY: CONTAINS 3 F5/8 TYPE A DOMAINS; EACH IS COMPOSED OF  
 CC 2 PLASTOCYANIN-LIKE REPEATS.  
 CC -1- SIMILARITY: STRONG, TO COAGULATION FACTOR V.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: L05573; AAA37385.1; -  
 DR PIR: A47004; A47004.  
 DR HSSP: P00451; ICFG.  
 DR MGD: MGI:88383; F8.  
 DR InterPro: IPR000421; -  
 DR InterPro: IPR001117; -  
 DR Pfam: PF00754; Cu-oxidase; 3.  
 DR Pfam: PF00394; F5-F8-type-C; 2.  
 DR PROSITE: PS00079; MULTICOPPER\_OXIDASE1; 3.  
 DR PROSITE: PS01285; FA58C\_1; 2.  
 DR PROSITE: PS01286; FA58C\_2; 2.  
 KM Blood coagulation: Repeat; Plasma; Acute phase; Calcium;  
 KW Signal: Glycoprotein; Sulfatation.  
 FT SIGNAL 1 19  
 FT CHAIN 20 2319  
 FT DOMAIN 20 349  
 FT DOMAIN 20 199  
 FT DOMAIN 207 349  
 FT DOMAIN 399 730  
 FT DOMAIN 399 573  
 FT DOMAIN 583 730  
 FT DOMAIN 760 1640  
 FT DOMAIN 1683 2008  
 FT DOMAIN 1683 1845  
 FT DOMAIN 1855 2008  
 FT DOMAIN 2008 2156  
 FT DOMAIN 2161 2313  
 FT SITE 391 392  
 FT SITE 391 392  
 FT SITE 759 760  
 FT SITE 1678 1679  
 FT SITE 1324 1325  
 FT SITE 1640 1641  
 FT MOD\_RES 367 367  
 FT MOD\_RES 367 367  
 FT MOD\_RES 738 738  
 FT MOD\_RES 742 742  
 FT MOD\_RES 1669 1669  
 FT MOD\_RES 1687 1687  
 FT DISULFID 173 199  
 FT DISULFID 547 573  
 FT DISULFID 1819 1845  
 FT DISULFID 2008 2156  
 FT DISULFID 2161 2313

FT CARBOHYD 61 61 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 233 233 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 259 259 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 423 423 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 601 601 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 880 880 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 958 958 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1015 1015 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1022 1022 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1026 1026 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1044 1044 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1076 1076 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1087 1087 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1136 1136 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1161 1161 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1192 1192 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1255 1255 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1268 1268 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1273 1273 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1274 1274 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1302 1302 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1316 1316 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1340 1340 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1378 1378 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 1797 1797 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 2105 2105 N-LINKED (GLCNAC...) (POTENTIAL).  
 SQ SEQUENCE 2319 AA; 266148 MW; FD054DE051DB2A01 CRC64;

Query Match 80.1%; Score 6301; DB 1; Length 2319;  
 Best Local Similarity 53.8%; Pred. No. 0;  
 Matches 1251; Conservative 102; Mismatches 111; Indels 862; Gaps 10;

QY 1 MDELSTCEFLCLRCFCFSATRRYYLGAVELSMQWQSD-LGELPVDAFRPPRVKSPFE 59  
 DB 1 MQALFACFFLLSFNFCSSAIRYYLGAVELSMNIIQSLLVLTDSRFLPRMSTSPFE 60  
 QY 60 NTSVYKKTLEVEFVHLEINAKPRPMMGLIPTQAEVDYTVVITLKNNASHPVSJHA 119  
 DB 61 NTSIMKKTIVFVEYKQDLNIAKPRPMMGLIPTQAEVDYTVVITLKNNASHPVSJHA 120  
 QY 120 VGVSYKASGEAYDQGTQREKEDKVPFGSHYYWQVLEKNPMPASDPLCTYSLS 179  
 DB 121 VGVSYKASGEAYDQGTQREKEDKVPFGSHYYWQVLEKNPMPASDPLCTYSLS 180  
 QY 180 HDVLVDLNSGLIGALLVCREGSLAKETQTLHKFTLFAVDEKSMHSEKNSLMODR 239  
 DB 181 HDVLVDLNSGLIGALLVCREGSLAKETQTLHKFTLFAVDEKSMHSEKNSLMODR 240  
 QY 240 DAASARAMPKMTVNGVYVNRSLPGLIGCHRSYVHVHVGMTTPVHSTFLEGHTELYRN 299  
 DB 241 DSASARDMPKMTVNGVYVNRSLPGLIGCHRSYVHVHVGMTTPVHSTFLEGHTELYRN 300  
 QY 300 HRASLEISPIFTFLAQTLLMLGOLFCHTSSHOHOMEMAYVVDSCPREPOLRMK-N 358  
 DB 301 HRASLEISPIFTFLAQTLLMLGOLFCHTSSHOHOMEMAYVVDSCPREPOLRMK-N 360  
 QY 359 NEAEADYDDELDEMDVVRFPDDNSPFIQIRSAKRPKRVYVIAAEEDMDYALV 418  
 DB 361 NEMEDYDDELDEMDVVRFPDDNSPFIQIRSAKRPKRVYVIAAEEDMDYALV 418  
 QY 419 LAPDRSYKQYLLNNGPQRIQIRKRVPRMAYTDEFKTRIAIQHESGILGLAYGEVD 478  
 DB 419 PYSNDGYSQYLLNNGPQRIQIRKRVPRMAYTDEFKTRIAIQHESGILGLAYGEVD 478  
 QY 479 TLLIFKNOASRPYNIYPGITDVPRPLYSRKLPGKYKHLKDFLLPGELFKKTVYED 538  
 DB 479 TLLIFKNOASRPYNIYPGITDVPRPLYSRKLPGKYKHLKDFLLPGELFKKTVYED 538  
 QY 539 GPTKSPDCLTRYSSFVMMERDLASGLIGPLLYCKESVDORGNOIMSDRNVTLFVSF 598  
 DB 539 GPTKSPDCLTRYSSFVMMERDLASGLIGPLLYCKESVDORGNOIMSDRNVTLFVSF 598



RA Kane W.H., Ichinose A., Hagen F.S., Davie E.W.;  
 RT "Cloning of cDNAs coding for the heavy chain region and connecting  
 RT region of human factor V, a blood coagulation factor with four types  
 RT of internal repeats.";  
 RL Biochemistry 26:6508-6514(1987).  
 RN [4]  
 RP SEQUENCE OF 1188-1215 AND 1315-2224 FROM N.A.  
 RX MEDLINE=86313665; PubMed=30922220;  
 RA Kane W.H., Davie E.W.;  
 RT "Cloning of a cDNA coding for human factor V, a blood coagulation  
 RT factor homologous to factor VIII and ceruloplasmin.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:6800-6804(1986).  
 RN [5]  
 RP PARTIAL SEQUENCE FROM N.A.  
 RC TISSUE-Fibrinolytic;  
 RX MEDLINE=93203619; PubMed=8454869;  
 RA Shen N.L., Fan S.-T., Pyatt J., Graff R., Lapolla R.J.,  
 RA Edgington T.S.;  
 RT "The serine protease cofactor factor V is synthesized by  
 RT lymphocytes.";  
 RL J. Immunol. 150:2992-3001(1993).  
 RN [6]  
 RP VARIANT APCR GLN-534.  
 RX MEDLINE=94217810; PubMed=8164741;  
 RA Bertina R.M., Koelman B.P.C., Koster T., Rosendaal F.R.,  
 RA Driven R.J., de Ronde H., van der Velden P.A., Reitsma P.H.;  
 RT "Mutation in blood coagulation factor V associated with resistance to  
 RT activated protein C.";  
 RL Nature 369:64-67(1994).  
 CC -1- FUNCTION: COAGULATION FACTOR V IS A COFACTOR THAT PARTICIPATES  
 CC WITH FACTOR XA TO ACTIVATE PROTHROMBIN TO THROMBIN.  
 CC -1- SUBUNIT: FACTOR VA IS COMPOSED OF AN HEAVY CHAIN AND OF A LIGHT  
 CC CHAIN NONCOVALENTLY BOUND. THE INTERACTION BETWEEN THE TWO CHAINS  
 CC IS CALCIUM-DEPENDENT.  
 CC -1- DOMAIN: DOMAIN B CONTAINS 35 X 9 AA TANDEM REPEATS, AND 2 X 17 AA  
 CC REPEATS.  
 CC -1- PRM: THROMBIN ACTIVATES FACTOR V PROTEOLYTICALLY TO THE ACTIVE  
 CC COFACTOR, FACTOR V(A) (FORMATION OF A HEAVY CHAIN AT THE N-  
 CC TERMINUS AND A LIGHT CHAIN AT THE C-TERMINUS).  
 CC -1- DISEASE: OMEN PARAHEMOPHILIA, AN HEMORRHAGIC DIATHESIS, IS DUE  
 CC TO A DEFICIENCY OF FACTOR V. OTHER DEFECTS IN F5 RESULTS IN A  
 CC FORM OF THROMBOPHILIA KNOWN AS APC RESISTANCE (APCR). THE APCR  
 CC MUTATION IS FOUND IN ABOUT 5% OF THE POPULATION WHICH SUGGESTS THAT  
 CC A SLIGHT THROMBOTIC TENDENCY MAY CONFER SOME ADVANTAGE IN FETAL  
 CC IMPLANTATION.  
 CC -1- SIMILARITY: CONTAINS 3 F5/8 TYPE A DOMAINS; EACH IS COMPOSED OF  
 CC 2 PLASTOCYANIN-LIKE REPEATS.  
 CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.  
 CC -1- SIMILARITY: STRONG, TO COAGULATION FACTOR VIII.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sb.ch/announce/>  
 CC or send an email to [license@isb-sb.ch](mailto:license@isb-sb.ch)).  
 CC -----  
 DR EMBL; L32779; AAB59401.1; JOINED.  
 DR EMBL; L32755; AAB59401.1; JOINED.  
 DR EMBL; L32756; AAB59401.1; JOINED.  
 DR EMBL; L32757; AAB59401.1; JOINED.  
 DR EMBL; L32758; AAB59401.1; JOINED.  
 DR EMBL; L32759; AAB59401.1; JOINED.  
 DR EMBL; L32760; AAB59401.1; JOINED.  
 DR EMBL; L32761; AAB59401.1; JOINED.  
 DR EMBL; L32762; AAB59401.1; JOINED.  
 DR EMBL; L32763; AAB59401.1; JOINED.  
 DR EMBL; L32764; AAB59401.1; JOINED.  
 DR EMBL; L32765; AAB59401.1; JOINED.  
 DR EMBL; L32766; AAB59401.1; JOINED.  
 DR EMBL; L32767; AAB59401.1; JOINED.  
 DR EMBL; L32768; AAB59401.1; JOINED.

DR EMBL; L32769; AAB59401.1; JOINED.  
 DR EMBL; L32770; AAB59401.1; JOINED.  
 DR EMBL; L32771; AAB59401.1; JOINED.  
 DR EMBL; L32772; AAB59401.1; JOINED.  
 DR EMBL; L32773; AAB59401.1; JOINED.  
 DR EMBL; L32774; AAB59401.1; JOINED.  
 DR EMBL; L32775; AAB59401.1; JOINED.  
 DR EMBL; L32776; AAB59401.1; JOINED.  
 DR EMBL; L32777; AAB59401.1; JOINED.  
 DR EMBL; L32778; AAB59401.1; JOINED.  
 DR EMBL; L32779; AAB59401.1; JOINED.  
 DR EMBL; L32780; AAB59401.1; JOINED.  
 DR EMBL; L32781; AAB59401.1; JOINED.  
 DR EMBL; L32782; AAB59401.1; JOINED.  
 DR EMBL; L32783; AAB59401.1; JOINED.  
 DR EMBL; L32784; AAB59401.1; JOINED.  
 DR EMBL; L32785; AAB59401.1; JOINED.  
 DR EMBL; L32786; AAB59401.1; JOINED.  
 DR EMBL; L32787; AAB59401.1; JOINED.  
 DR EMBL; L32788; AAB59401.1; JOINED.  
 DR EMBL; L32789; AAB59401.1; JOINED.  
 DR EMBL; L32790; AAB59401.1; JOINED.  
 DR EMBL; L32791; AAB59401.1; JOINED.  
 DR EMBL; L32792; AAB59401.1; JOINED.  
 DR EMBL; L32793; AAB59401.1; JOINED.  
 DR EMBL; L32794; AAB59401.1; JOINED.  
 DR EMBL; L32795; AAB59401.1; JOINED.  
 DR EMBL; L32796; AAB59401.1; JOINED.  
 DR EMBL; L32797; AAB59401.1; JOINED.  
 DR EMBL; L32798; AAB59401.1; JOINED.  
 DR EMBL; L32799; AAB59401.1; JOINED.  
 DR EMBL; L32800; AAB59401.1; JOINED.  
 DR EMBL; L32801; AAB59401.1; JOINED.  
 DR EMBL; L32802; AAB59401.1; JOINED.  
 DR EMBL; L32803; AAB59401.1; JOINED.  
 DR EMBL; L32804; AAB59401.1; JOINED.  
 DR EMBL; L32805; AAB59401.1; JOINED.  
 DR EMBL; L32806; AAB59401.1; JOINED.  
 DR EMBL; L32807; AAB59401.1; JOINED.  
 DR EMBL; L32808; AAB59401.1; JOINED.  
 DR EMBL; L32809; AAB59401.1; JOINED.  
 DR EMBL; L32810; AAB59401.1; JOINED.  
 DR EMBL; L32811; AAB59401.1; JOINED.  
 DR EMBL; L32812; AAB59401.1; JOINED.  
 DR EMBL; L32813; AAB59401.1; JOINED.  
 DR EMBL; L32814; AAB59401.1; JOINED.  
 DR EMBL; L32815; AAB59401.1; JOINED.  
 DR EMBL; L32816; AAB59401.1; JOINED.  
 DR EMBL; L32817; AAB59401.1; JOINED.  
 DR EMBL; L32818; AAB59401.1; JOINED.  
 DR EMBL; L32819; AAB59401.1; JOINED.  
 DR EMBL; L32820; AAB59401.1; JOINED.  
 DR EMBL; L32821; AAB59401.1; JOINED.  
 DR EMBL; L32822; AAB59401.1; JOINED.  
 DR EMBL; L32823; AAB59401.1; JOINED.  
 DR EMBL; L32824; AAB59401.1; JOINED.  
 DR EMBL; L32825; AAB59401.1; JOINED.  
 DR EMBL; L32826; AAB59401.1; JOINED.  
 DR EMBL; L32827; AAB59401.1; JOINED.  
 DR EMBL; L32828; AAB59401.1; JOINED.  
 DR EMBL; L32829; AAB59401.1; JOINED.  
 DR EMBL; L32830; AAB59401.1; JOINED.  
 DR EMBL; L32831; AAB59401.1; JOINED.  
 DR EMBL; L32832; AAB59401.1; JOINED.  
 DR EMBL; L32833; AAB59401.1; JOINED.  
 DR EMBL; L32834; AAB59401.1; JOINED.  
 DR EMBL; L32835; AAB59401.1; JOINED.  
 DR EMBL; L32836; AAB59401.1; JOINED.  
 DR EMBL; L32837; AAB59401.1; JOINED.  
 DR EMBL; L32838; AAB59401.1; JOINED.  
 DR EMBL; L32839; AAB59401.1; JOINED.  
 DR EMBL; L32840; AAB59401.1; JOINED.  
 DR EMBL; L32841; AAB59401.1; JOINED.  
 DR EMBL; L32842; AAB59401.1; JOINED.  
 DR EMBL; L32843; AAB59401.1; JOINED.  
 DR EMBL; L32844; AAB59401.1; JOINED.  
 DR EMBL; L32845; AAB59401.1; JOINED.  
 DR EMBL; L32846; AAB59401.1; JOINED.  
 DR EMBL; L32847; AAB59401.1; JOINED.  
 DR EMBL; L32848; AAB59401.1; JOINED.  
 DR EMBL; L32849; AAB59401.1; JOINED.  
 DR EMBL; L32850; AAB59401.1; JOINED.  
 DR EMBL; L32851; AAB59401.1; JOINED.  
 DR EMBL; L32852; AAB59401.1; JOINED.  
 DR EMBL; L32853; AAB59401.1; JOINED.  
 DR EMBL; L32854; AAB59401.1; JOINED.  
 DR EMBL; L32855; AAB59401.1; JOINED.  
 DR EMBL; L32856; AAB59401.1; JOINED.  
 DR EMBL; L32857; AAB59401.1; JOINED.  
 DR EMBL; L32858; AAB59401.1; JOINED.  
 DR EMBL; L32859; AAB59401.1; JOINED.  
 DR EMBL; L32860; AAB59401.1; JOINED.  
 DR EMBL; L32861; AAB59401.1; JOINED.  
 DR EMBL; L32862; AAB59401.1; JOINED.  
 DR EMBL; L32863; AAB59401.1; JOINED.  
 DR EMBL; L32864; AAB59401.1; JOINED.  
 DR EMBL; L32865; AAB59401.1; JOINED.  
 DR EMBL; L32866; AAB59401.1; JOINED.  
 DR EMBL; L32867; AAB59401.1; JOINED.  
 DR EMBL; L32868; AAB59401.1; JOINED.  
 DR EMBL; L32869; AAB59401.1; JOINED.  
 DR EMBL; L32870; AAB59401.1; JOINED.  
 DR EMBL; L32871; AAB59401.1; JOINED.  
 DR EMBL; L32872; AAB59401.1; JOINED.  
 DR EMBL; L32873; AAB59401.1; JOINED.  
 DR EMBL; L32874; AAB59401.1; JOINED.  
 DR EMBL; L32875; AAB59401.1; JOINED.  
 DR EMBL; L32876; AAB59401.1; JOINED.  
 DR EMBL; L32877; AAB59401.1; JOINED.  
 DR EMBL; L32878; AAB59401.1; JOINED.  
 DR EMBL; L32879; AAB59401.1; JOINED.  
 DR EMBL; L32880; AAB59401.1; JOINED.  
 DR EMBL; L32881; AAB59401.1; JOINED.  
 DR EMBL; L32882; AAB59401.1; JOINED.  
 DR EMBL; L32883; AAB59401.1; JOINED.  
 DR EMBL; L32884; AAB59401.1; JOINED.  
 DR EMBL; L32885; AAB59401.1; JOINED.  
 DR EMBL; L32886; AAB59401.1; JOINED.  
 DR EMBL; L32887; AAB59401.1; JOINED.  
 DR EMBL; L32888; AAB59401.1; JOINED.  
 DR EMBL; L32889; AAB59401.1; JOINED.  
 DR EMBL; L32890; AAB59401.1; JOINED.  
 DR EMBL; L32891; AAB59401.1; JOINED.  
 DR EMBL; L32892; AAB59401.1; JOINED.  
 DR EMBL; L32893; AAB59401.1; JOINED.  
 DR EMBL; L32894; AAB59401.1; JOINED.  
 DR EMBL; L32895; AAB59401.1; JOINED.  
 DR EMBL; L32896; AAB59401.1; JOINED.  
 DR EMBL; L32897; AAB59401.1; JOINED.  
 DR EMBL; L32898; AAB59401.1; JOINED.  
 DR EMBL; L32899; AAB59401.1; JOINED.  
 DR EMBL; L32900; AAB59401.1; JOINED.  
 DR EMBL; L32901; AAB59401.1; JOINED.  
 DR EMBL; L32902; AAB59401.1; JOINED.  
 DR EMBL; L32903; AAB59401.1; JOINED.  
 DR EMBL; L32904; AAB59401.1; JOINED.  
 DR EMBL; L32905; AAB59401.1; JOINED.  
 DR EMBL; L32906; AAB59401.1; JOINED.  
 DR EMBL; L32907; AAB59401.1; JOINED.  
 DR EMBL; L32908; AAB59401.1; JOINED.  
 DR EMBL; L32909; AAB59401.1; JOINED.  
 DR EMBL; L32910; AAB59401.1; JOINED.  
 DR EMBL; L32911; AAB59401.1; JOINED.  
 DR EMBL; L32912; AAB59401.1; JOINED.  
 DR EMBL; L32913; AAB59401.1; JOINED.  
 DR EMBL; L32914; AAB59401.1; JOINED.  
 DR EMBL; L32915; AAB59401.1; JOINED.  
 DR EMBL; L32916; AAB59401.1; JOINED.  
 DR EMBL; L32917; AAB59401.1; JOINED.  
 DR EMBL; L32918; AAB59401.1; JOINED.  
 DR EMBL; L32919; AAB59401.1; JOINED.  
 DR EMBL; L32920; AAB59401.1; JOINED.  
 DR EMBL; L32921; AAB59401.1; JOINED.  
 DR EMBL; L32922; AAB59401.1; JOINED.  
 DR EMBL; L32923; AAB59401.1; JOINED.  
 DR EMBL; L32924; AAB59401.1; JOINED.  
 DR EMBL; L32925; AAB59401.1; JOINED.  
 DR EMBL; L32926; AAB59401.1; JOINED.  
 DR EMBL; L32927; AAB59401.1; JOINED.  
 DR EMBL; L32928; AAB59401.1; JOINED.  
 DR EMBL; L32929; AAB59401.1; JOINED.  
 DR EMBL; L32930; AAB59401.1; JOINED.  
 DR EMBL; L32931; AAB59401.1; JOINED.  
 DR EMBL; L32932; AAB59401.1; JOINED.  
 DR EMBL; L32933; AAB59401.1; JOINED.  
 DR EMBL; L32934; AAB59401.1; JOINED.  
 DR EMBL; L32935; AAB59401.1; JOINED.  
 DR EMBL; L32936; AAB59401.1; JOINED.  
 DR EMBL; L32937; AAB59401.1; JOINED.  
 DR EMBL; L32938; AAB59401.1; JOINED.  
 DR EMBL; L32939; AAB59401.1; JOINED.  
 DR EMBL; L32940; AAB59401.1; JOINED.  
 DR EMBL; L32941; AAB59401.1; JOINED.  
 DR EMBL; L32942; AAB59401.1; JOINED.  
 DR EMBL; L32943; AAB59401.1; JOINED.  
 DR EMBL; L32944; AAB59401.1; JOINED.  
 DR EMBL; L32945; AAB59401.1; JOINED.  
 DR EMBL; L32946; AAB59401.1; JOINED.  
 DR EMBL; L32947; AAB59401.1; JOINED.  
 DR EMBL; L32948; AAB59401.1; JOINED.  
 DR EMBL; L32949; AAB59401.1; JOINED.  
 DR EMBL; L32950; AAB59401.1; JOINED.  
 DR EMBL; L32951; AAB59401.1; JOINED.  
 DR EMBL; L32952; AAB59401.1; JOINED.  
 DR EMBL; L32953; AAB59401.1; JOINED.  
 DR EMBL; L32954; AAB59401.1; JOINED.  
 DR EMBL; L32955; AAB59401.1; JOINED.  
 DR EMBL; L32956; AAB59401.1; JOINED.  
 DR EMBL; L32957; AAB59401.1; JOINED.  
 DR EMBL; L32958; AAB59401.1; JOINED.  
 DR EMBL; L32959; AAB59401.1; JOINED.  
 DR EMBL; L32960; AAB59401.1; JOINED.  
 DR EMBL; L32961; AAB59401.1; JOINED.  
 DR EMBL; L32962; AAB59401.1; JOINED.  
 DR EMBL; L32963; AAB59401.1; JOINED.  
 DR EMBL; L32964; AAB59401.1; JOINED.  
 DR EMBL; L32965; AAB59401.1; JOINED.  
 DR EMBL; L32966; AAB59401.1; JOINED.  
 DR EMBL; L32967; AAB59401.1; JOINED.  
 DR EMBL; L32968; AAB59401.1; JOINED.  
 DR EMBL; L32969; AAB59401.1; JOINED.  
 DR EMBL; L32970; AAB59401.1; JOINED.  
 DR EMBL; L32971; AAB59401.1; JOINED.  
 DR EMBL; L32972; AAB59401.1; JOINED.  
 DR EMBL; L32973; AAB59401.1; JOINED.  
 DR EMBL; L32974; AAB59401.1; JOINED.  
 DR EMBL; L32975; AAB59401.1; JOINED.  
 DR EMBL; L32976; AAB59401.1; JOINED.  
 DR EMBL; L32977; AAB59401.1; JOINED.  
 DR EMBL; L32978; AAB59401.1; JOINED.  
 DR EMBL; L32979; AAB59401.1; JOINED.  
 DR EMBL; L32980; AAB59401.1; JOINED.  
 DR EMBL; L32981; AAB59401.1; JOINED.  
 DR EMBL; L32982; AAB59401.1; JOINED.  
 DR EMBL; L32983; AAB59401.1; JOINED.  
 DR EMBL; L32984; AAB59401.1; JOINED.  
 DR EMBL; L32985; AAB59401.1; JOINED.  
 DR EMBL; L32986; AAB59401.1; JOINED.  
 DR EMBL; L32987; AAB59401.1; JOINED.  
 DR EMBL; L32988; AAB59401.1; JOINED.  
 DR EMBL; L32989; AAB59401.1; JOINED.  
 DR EMBL; L32990; AAB59401.1; JOINED.  
 DR EMBL; L32991; AAB59401.1; JOINED.  
 DR EMBL; L32992; AAB59401.1; JOINED.  
 DR EMBL; L32993; AAB59401.1; JOINED.  
 DR EMBL; L32994; AAB59401.1; JOINED.  
 DR EMBL; L32995; AAB59401.1; JOINED.  
 DR EMBL; L32996; AAB59401.1; JOINED.  
 DR EMBL; L32997; AAB59401.1; JOINED.  
 DR EMBL; L32998; AAB59401.1; JOINED.  
 DR EMBL; L32999; AAB59401.1; JOINED.  
 DR EMBL; L33000; AAB59401.1; JOINED.

```

FT REPEAT 1428 1436 28.
FT REPEAT 1437 1445 29.
FT REPEAT 1446 1454 30.
FT REPEAT 1455 1463 31.
FT REPEAT 1464 1472 32.
FT REPEAT 1473 1481 33.
FT REPEAT 1482 1490 34.
FT REPEAT 1493 1501 35.
FT DOMAIN 1578 1597 F5/8 TYPE A 3.
FT DOMAIN 1578 1597 PLASTOCYANIN-LIKE 5.
FT DOMAIN 1761 1907 PLASTOCYANIN-LIKE 6.
FT DOMAIN 1907 2061 F5/8 TYPE C 1.
FT SITE 2066 2221 F5/8 TYPE C 2.
FT SITE 737 738 CLEAVAGE (BY THROMBIN).
FT SITE 1046 1047 CLEAVAGE (BY THROMBIN).
FT SITE 1573 1574 CLEAVAGE (BY THROMBIN).
FT DISULFID 167 193 PROBABLE.
FT DISULFID 500 526 PROBABLE.
FT DISULFID 1725 1751 PROBABLE.
FT DISULFID 1907 2061 BY SIMILARITY.
FT DISULFID 2066 2221 BY SIMILARITY.
FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 55 55 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 460 460 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

Query Match 30.6%, Score 2405.5; DB 1; Length 2224;  
 Best Local Similarity 26.3%; Pred. No. 3e-149;  
 Matches 599; Conservative 278; Mismatches 490; Indels 909; Gaps 37;

```

OY 22 RRYLGAIVELSDMYDQSLGELPVDARPPRRPKSPFNTSYV-YKRLVFEYTHLENI 80
DB 32 RQFYVAAGIGISYRPE-----PTNSSLNLSTVSKKLYREPEY-FKK 75
OY 81 AKRPPMMGLLPTIOAEVYDTVTLLKNMASHPYSLHAIVGSYMKASBGAEYDQTSOR 140
DB 76 EKRPQSTIGLGLPTIAXEYGDILIKYFNKAKPLSTIHQIGIRISKLSGASYLDHTPPA 135
OY 141 EKEDKVPFGSGSHYVWOLKENGPMASDPLCLVSYLSHVDLVKDLNSGLIGALLVGRE 200
DB 136 EKMDAIVAPGREYEWSEISEDSGPTHDDPCLTHIYSHENLIDPNSGLIGLPLICK 195
OY 201 GSIAEKQ-Q-ILHK-FILLPAVFDGSKMHSSTKNSLMODRDAASARAPPKMTYNGYVN 258
DB 196 GILTEGGTQKTFDKQIVLLFAVFDKSKMSOSS-----LMYTYNGYVN 239
OY 259 RSLPLGICHRKSYWYHIGMGTPEVHSIFLEGHTFLVRNHRQASLEISPTFLAQT 318
DB 240 GIMPDTIYCAHDHISNHLIGMSSGPELFSIHNGOYLEQNHKYSATITLVASTITANMT 299
OY 319 LMDLGQFLFCHISSHODGMEAYVYKVDSCPEEQPLRMKNNEAEADYDDDLTDEMDYVR 378
DB 300 VQPEGKWIISLTPRKHLOAGMOAYIDIKNCPKRTNKLKTRP----- 342
OY 379 FDDDSPFQIRSVAKKHPTWYHIAEEDMDYAPLVAPDRSTKSYQLANGPQRT 438
DB 343 -----QRHMKRMEYFIAAEVIMDYAPVPIRANMKKRSQSLDFNSQI 387
OY 439 GRKYKVFMAVYDTEFTTREAI---QHESGILGPLLGEYDGLTLLIFKNQASRPYNY 495
DB 388 GKHYKVMYTOYEDESF-TKHTVNPKNKEDGILGPIIAQYRDTLKIYFKMNASPYSIY 446
OY 496 PHGIT-----DVRPLYSRRLPKGVKHLKDRPILPGLFIKRYKTVYVDEGFTKSDPRCLT 549
DB 447 PHGVTFSEYDEVNSSF-----SGRNTMTIRAYVQGETYTYKWNILLEDEPTENDAOGLT 502
OY 550 KYSSVFVMERDLASGLIGPLLIKESYDQNGOIMDKRNVILFSYFENRSMYLTEN 609
DB 503 RPYSDVADIMBIDIASGLIGLLICKSRSLDRGIGORADICQAAVFAVDEKSKSYLSDN 562
OY 610 IORFLPNAGVQOLDEPFGQASINMHSINGYVDSL-QLSVCLHEVAYVYLLISGQOTFL 668

```

```

DB 563 IKPCENPEDEVARRDPKFEYSINMSTINGYVDESITITLGFCHDDPQVMHFGSVQONEL 622
OY 669 SVFSGYTFKKHMYEDITLTPFSGEYVFMENPGLMILGCHNS----- 714
DB 623 THTFGHSFYGKRREHDTLLTPMRGESVYTMNVGMTLSMNSSPRSKRLRLKFRDY 682
OY 715 -----DFRNR----- 719
DB 683 KCIPIDDEDSYEIEFPPESTVATRKMDRLPEDEBESDADYQONRLAALGISFRNS 742
OY 720 -----GMTAL----- 724
DB 743 SLNQEERENLALLENGETEVSSNTDIIVGSNYSPPENISKFTYNNLAEPQKARSHQ 802
OY 725 -----LKYSCDKNTGDIYEDSYD----- 744
DB 803 ATTAGSPRLHLIGKNSVLSSTAHSHPSEDPIDPQDPYTGIRLLSLGAEFGSOEH 862
OY 745 ----- 744
DB 863 AKHKGPKYERDQAAKHRSWMLLAHKVGRHLSQDTGSPSGMRPWEDLPQDTGSPSRMR 922
OY 745 -----ISAVLSKRNA----- 755
DB 923 PKKDPSPDLLLKQSNSSKILVGRHHLASEKSGYELIQDTEDTAVNWLISPOAASRAW 982
OY 756 ----- 755
DB 983 GSETPLANKPGKSGHPKFPVRNHRSLQVRDGGKSRLLKSQFLIKTRKKKKKHTHNAF 1042
OY 756 IEPRSF-----SQ 763
DB 1043 LSPKTFHPLRSEAYVTFSEERRLKSLVLKSNETSPLTDNLQTLPSMDFGWIASLPDHQ 1102
OY 764 NSRH-----PSTR----- 771
DB 1103 NSSNPTGASCPGLXQTVPREHQTPIQDPDMHSTSPSHSSSELSMELEYDRS 1162
OY 772 OKQF-----NATPYLKRH- 785
DB 1163 HKSFTDISQMSSESEHEWQTVISPDLSQVTLSPELSQTNLSPDLSHTTSLPESLQRL 1222
OY 786 ----- 785
DB 1223 SPALGQMPISPDLSHTTSLSPDLSHTTSLDLSQTNLSPELSQTNLSPALGQMPISPDLSH 1282
OY 786 ----- 785
DB 1283 TTLSDPFSQTNLSPELSHMTLSPELSQTNLSPALGQMPISPDLSHTTSLDPSQTNLSPE 1342
OY 786 -----QREITRTTLQSDQEI- 801
DB 1343 LSQTNLSPALGQMPISPDPSHTTSLDLSQTNLSPELSQTNLSPDLSMPPLPADLSQIPL 1402
OY 802 ----- 801
DB 1403 TPDLQMTLSPDLGETDLPNFGQMSLSPDLSQVTLSPDISDTLLPDLSPQISPPDLQ 1462
OY 802 -----DYDT 806
DB 1463 IFYSESSQSLLOEFNESPYPDLGQMPSPSSPPLNDTFLSKERNPLVIYGLSKSGIDY 1522
OY 807 ISVENKKE-----DFDIYDE-----DENQSP-----RSFOKKTBY 837
DB 1523 IETIPKEVQSSQEDYATIDVYPYDDPKYTDVRTINSSRPDNIAMYLNNNNRNNY 1582
OY 838 FIAAVERLMDYGMSSSPYVLRNR-AQSGSVQ--FKKVVFOETDGSFTQPLKGLN 892
DB 1583 YIAAEISWDY-----SEVQRETDIEDSDDIPEDTYYKVVFRKLYDSTFKRPRGEYE 1638
OY 893 EHLGLGPIYIAEVEDNIMTFERNQASRPYSYSLIYE-----EDRQGAERPKN 944

```



```

FT DOMAIN 1748 1890 PLASTOCYANIN-LIKE 6.
FT DOMAIN 1894 2048 F5/8 TYPE C 1.
FT DOMAIN 2053 2208 F5/8 TYPE C 2.
FT SITE 741 742 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT SITE 1034 1035 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT SITE 1564 1565 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT DISULFID 167 193 PROBABLE.
FT DISULFID 499 525 PROBABLE.
FT DISULFID 1712 1738 PROBABLE.
FT DISULFID 1894 2048 BY SIMILARITY.
FT DISULFID 2053 2208 BY SIMILARITY.
FT CARBOHYD 225 225 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 460 460 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 553 553 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 587 587 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 745 745 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 756 756 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 774 774 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 780 780 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 902 902 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 952 952 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 964 964 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1044 1044 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1053 1053 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1062 1062 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1071 1071 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1078 1078 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1094 1094 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1451 1451 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1490 1490 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1550 1550 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1690 1690 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1839 1839 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1997 1997 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2196 2196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 587 592 NPTLPA -> T (IN VARIANT 2).
SQ SEQUENCE 2211 AA: 248981 MW: CBBF90B738667C45 CRC64:
```

Query Match 30.2%; Score 2379; DB 1; Length 2211;  
Best Local Similarity 26.3%; Pred. No. 1.6e-147;  
Matches 597; Conservative 267; Mismatches 496; Indels 910; Gaps 37;

```

QY 22 RRYVIGAVELSDVYMQSLGELPVDAREPPRYPKSPENTSVYKTLFEVETVHLENIA 81
DB 32 RQFYVAASQIRWNYR-----PESTHL-----SSKPFETS--FKIYVREYAV-FOKE 76
QY 82 KRPPMGLGPTIOAEVYDTVITLKNASHPVSLHANGVSYWKASGEAEVDDOTSORE 141
DB 77 KFOSTSGLEGTLYAEVGDILKVFHKNAKPLSIHAGIKYKFSSEASYSDTLPEME 136
QY 142 KEDKVFPGSGHTYVQVYLKENGPMASDPLCTLYSYLSHVLDVKNLSGILGALLVCREG 201
DB 137 KMDDAVNAAGQETTYEMITSEHSGPETHDPCPLTHIYYSVNLVEDPNSGLIPPLICKKG 196
QY 202 SLAKKKTQTL--HKFILLPAVFDECKSWHSETKNSLMQDRDAASARAMPKMTGVGNVR 259
DB 197 TLTEGTGOKMEKQHVLMFAVEDEKSWNQTS-----LMTVNGCYING 240
QY 260 SLPGILGHRKSVYVHYVGMGTPEVHSTFLEGHFTLVNRHQASLESPITFLAQOLL 319
DB 241 TMDPDLTVCAHDIHSLWHLIGMSSGPPELFSIHENGQVLEONHRIKISAITVVASITJANMTV 300
QY 330 MDLGLFLFCHLISHQHDGMEAYVAVKSCPEEPOLMRKNNEEAEDYDDDLTSEMDVYVF 379
DB 301 SPEGWTIASLIPRHFGQGMQAVYIDIKKCAKTRNPK-----LTFMDQ----- 343
QY 360 DDDNSPSTIQIRSAKKPKPTVNHITAAEEDMDYAPLVLAADPDRSYSQYVYINNGPQIRG 439
DB 344 -----RRHIKWEYFIAAEEVIMDYAPIIIPAMDKKYSLSLHDFNSNRIG 388
```

```

QY 440 RRYKVRFMAYVDETEKTR--EAIOHESGILGLYGEVQDTLLIFKNQASRPYNIYH 497
DB 389 KHYKVVYQYODDSFTKREDDSSGEGDILGLIAQVYDRLKIFYKNNASYSIYH 448
QY 498 GIT-----DVRLYSRRLPKGVKHLKDPILPGEIEFKYKWTIVEDGPKSPDRL 548
DB 449 GVFESYDNEVNSSSTSGSNMTIRAVR-----PGEIYTYKNNLLEDEDEPENAQCL 500
QY 549 TRYSSFVNMERDLASGLIGPLILCYKESVDORGNQIMSDKRVILFSVDENRSMYLTE 608
DB 501 TRPYYSNVITRDLASGLIGLILICKRSIDRGIORADIEQAAVFAVEDEKSWIYED 560
QY 609 NIORFLPNAGVQLEPEPOASNIMH-----SINGVPSIQ--LSCLHEVAWYILLISG 662
DB 561 NIYKFEENPEKRRDPPKFEYSNIMSNFTLPALINGVPEIIPILGFCFDYTVQMHFCVSG 620
QY 663 AOTDFLVSFGYTFEKHKVYEDTLTLPFSGEYVMSMENPGLMILGCHNSDFRNKMT 722
DB 621 TQNDILTIHFTGHSFYGKRHEDTLTLFPMQGESYVTJMDNVGTWMLTTMNSPRSKLR 680
QY 723 ALIKVSCCKNTGD-----YEDS-----YEDISAVL--- 749
DB 681 LRFRDAKCIKRNDDDSYEIIEYPSGSTAMTKKHDSSEIENDADSDYQDELALILGL 740
QY 750 ----- 749
DB 741 RSFRNSSLOEKDELNLTALALEKDEEITPPSANRSLDSSSRSHVSLIAKNFAESIK 800
QY 750 -----LSKNNAIEP----- 758
DB 801 TLLHLEAPAAGSPLEHAGLDKKSALNAPPAHESSPSEDPREDHPLSDYTVGSLLPFGTG 860
QY 759 -----RSFSQ----- 763
DB 861 FKNRKPAKHQRFQVGRGAQAKHKFSQTRPPAKHTRTLQSODNSSSRMGPMWEDPSDLL 920
QY 764 ----- 763
DB 921 LQOKDPYKILNGBMHLVSEKSEYEIIOANENKTYKNKLPSQNDSSRTNGENIPFNKSHG 980
QY 764 -NSRHPS-----TROKOFNATP-----PVL 782
DB 981 KQSGHPTFLVTRRRKPLQODQDRNSLKESGLPLIRTRRRKKKEKRPXVHYNPLSPRSHPLR 1040
QY 783 -----KRHOREI-----TRTTLQSDQEEI-----DYDD----- 805
DB 1041 GCVNASFSDRRNHNSILLHASNETSLSIDLQTFPSMNLSLAASLPDHDQTSPPNDTSQT 1100
QY 806 -----TISVEMKKEDFDIYDE-----ENQSP----- 827
DB 1101 SSPDLYPTVSPSEHYQIFPIQDSPTHTSTAPSNKSPDPTHTSTAPSNKSPPTQSPQIP 1160
QY 828 ----- 827
DB 1161 NYDLNRAPIPTVQSIFPSELEWQATSLDSQPSISPDLGQALSPDPQOESLSPDL 1220
QY 828 ----- 827
DB 1221 GQTSPLDLSQESLSPDLGQTALSPDPQOESLSPDLGQTALSPDPQOESLSPDLGQTALS 1280
QY 828 ----- 827
DB 1281 PDPQOESLSPDLGQTALSPDLDSQESLSPDLGQTALSPDPQOESLSPDLGQTALSPDPQOE 1340
QY 828 ----- 827
DB 1341 SLSPDLGQTALSPDLGQESLSPDLGQTALSPDPQOESLSPDLGQTALSPDLGQESLSPDL 1400
QY 828 ----- 827
DB 1401 GQTALSPDLDSQESLSPDLGQTALSPDLDSQESLSPDLGQTALSPDLGQTALSPDLGQESL 1460
```

```

QY 828 ----- 827
Db 1461 OSLPIPEFGOTFPNADIGQMPSPDPSTLNTNFIPEBFNLVYVGLSRDDGDYEIIPRQ 1520
QY 828 ----- RSOCKTRHFIAVER 844
Db 1521 KEESSEEDYGEFEFVAYNDPYQDTRTDINSSRNPNDAWYLRSNTGNKRYYYIAEEI 1580
QY 845 LMDYSSSPHYLRNAAOGSS---VPO---FKKVFQFOEFGSFTOPLYGENEHLGL 898
Db 1501 SMDYS-----KTVQSDDDVDYVEDIVYKKVYFRKXIDSTFTKIDPGEVEEHLGITL 1631
QY 899 GPYIRAEVEDNIMVTFRNOASRPYSEYSSLSIYE-----EDORQAGAPRRNVEYKPN 950
Db 1632 GPVIRAEVDDVIOVRKKNLASRPYSLAHGLSYEKSECKTYEDDSDPEWEKEDNADIPNK 1691
QY 951 TKTYWVKVQHNAAPTKDEEDCKAMAFESVDLEKDVHSGLIGPLVCHTTLNPAHRQV 1010
Db 1692 TYTYVWHATTRSGPENPGSACRAMAYSAVNPBKDIHSGLIGPLLCRKGTDLKETMPV 1751
QY 1011 TVQEFALFTTIDETKSMYFTEEMERNCRAPCNIQMEDPFEKENYRPHAINGYIMDTLP 1070
Db 1752 DMREYVLEFMVDEDEKKSWMYDKKPTSMWRASS-----EVKSHERHAINGMATYN-LPG 1804
QY 1071 LVMAODQIRWYLLSGMSNENIHSHFSGHVTFRKKEEYKMLYMLPGVFETVEMLP 1130
Db 1805 LRMVQEWVRLHLNLGSGRDIHVHFHCGQTLLENCTQHQHGLVPLPGSFKLEKKAS 1864
QY 1131 KAGIRVRELLIGEHLAGSTLEFLVYSNKCQPTPLGMASGHIIROPITASQYQOMAPKLA 1190
Db 1865 KFGWMLDDEVEGDIQAGKQGTPLIVDRECKMPGLSTGLIASQIQOASFEWGWEPKLA 1924
QY 1191 RLHYSGSIAM-----STK-EPPSWIKVDLAPMIIGITIGAROKFSSLYISQITIMY 1244
Db 1925 RLNNGSGYAMWIAEKSTEFENPEPMIOVDMQKEVLLTGIOGAKHKLKRYTTEFCVAY 1984
QY 1245 SLDGKRMQTYRGNSTGTLMVFEQVNDSSGIKHNIENPPIIARYIRLPHYTSIRSTLRME 1304
Db 1985 SLDRKMRIFKGNSTRNVMYFGNSDASTIKENDIDPEVARYIRISPTSYNKPALRLE 2044
QY 1305 LMGCDLNSCMLPGMSKALISDAQITASSYFTMFEAT-WSPSKARLHLQGRSNMWRQVN 1363
Db 2045 LOGCEVNGSTPLGMSGKIENKQITASSFKKSMGWYMEPPLARLMAOGSRVNAWQAKAN 2104
QY 1364 NPEKMLQVDFOKTMKYTGVTQGVSKSLTSMYKEFLISSODGHOWTLFFONGKV--KV 1421
Db 2105 NNNQWLOQIDLKIKKTAIVTQGCCKLSSEMYKSYTIHSDGDTMKPYREKSSMWDKI 2164
QY 1422 FQGNDSFTPVVNSLDPRLITRYLRHPOSWQHIALRMEVLCEADODLY 1471
Db 2165 FEGNNNVNRGHVKNFNPITISRFIRIIPKTNOSIALRLLEFGC---DMY 2211

RESULT 6
CERU_HUMAN STANDARD: PRT: 1065 AA.
AC P00450; Q14063;
DT 21-JUL-1986 (Rel. 01, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CERULOPLASMIN PRECURSOR (EC 1.16.3.1) (FERROXIDASE).
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86259737; PubMed=2873574;
RA Koschinsky M.L., Funk W.D., Van Oost B.A., McGillivray R.T.A.;
RT "Complete cDNA sequence of human preceruloplasmin.",
RL Proc. Natl. Acad. Sci. U.S.A. 83:5086-5090(1986).
[2]

```

```

RP SEQUENCE OF 1-1006 FROM N.A.
RX MEDLINE=95217183; PubMed=7702601;
RA Daimon M., Yamatani K., Igarashi M., Fukase N., Kawanami T.,
RA Kato T., Tomioka M., Sasaki H.;
RT "Fine structure of human ceruloplasmin gene.",
RL Biochem. Biophys. Res. Commun. 208:1028-1035(1995).
RN [3]
RP SEQUENCE OF 1-40; 549-599; 784-829 AND 919-952 FROM N.A.
RX MEDLINE=86275241; PubMed=3755405;
RA Mercer J.F.B., Grimes A.;
RT "Isolation of a human ceruloplasmin cDNA clone that includes the N-
RT terminal leader sequence.",
RL FEBS Lett. 203:185-190(1986).
RN [4]
RP SEQUENCE OF 218-1065 FROM N.A.
RX MEDLINE=86205876; PubMed=3486416;
RA Yang F., Naylor S.L., Lum J.B., Cuthshaw S., McCombs J.L.,
RA Naberhaus K.H., McGill J.R., Adrian G.S., Moore C.M., Barnett D.R.,
RA Bowman B.H.;
RT "Characterization, mapping, and expression of the human ceruloplasmin
RT gene.",
RL Proc. Natl. Acad. Sci. U.S.A. 83:3257-3261(1986).
RN [5]
RP SEQUENCE OF 20-1065.
RX MEDLINE=84119493; PubMed=6582496;
RA Takahashi N., Ortel T.L., Putnam F.W.;
RT "Single-chain structure of human ceruloplasmin: the complete amino
RT acid sequence of the whole molecule.",
RL Proc. Natl. Acad. Sci. U.S.A. 81:390-394(1984).
RN [6]
RP SEQUENCE OF 158-333; 518-724 AND 858-1065.
RX MEDLINE=83117800; PubMed=6571985;
RA Takahashi N., Bauman R.A., Ortel T.L., Dwulet F.E., Wang C.-C.,
RA Putnam F.W.;
RT "Internal triplication in the structure of human ceruloplasmin.",
RL Proc. Natl. Acad. Sci. U.S.A. 80:115-119(1983).
RN [7]
RP SEQUENCE OF 501-905.
RX MEDLINE=81199407; PubMed=6940148;
RA Dwulet F.E., Putnam F.W.;
RT "Complete amino acid sequence of a 50,000-dalton fragment of human
RT ceruloplasmin.",
RL Proc. Natl. Acad. Sci. U.S.A. 78:790-794(1981).
RN [8]
RP SEQUENCE OF 907-1065.
RX MEDLINE=80137543; PubMed=6987229;
RA Kingston I.B., Kingston B.L., Putnam F.W.;
RT "Primary structure of a histidine-rich proteolytic fragment of human
RT ceruloplasmin. I. Amino acid sequence of the cyanogen bromide
RT peptides.",
RL J. Biol. Chem. 255:2878-2885(1980).
RN [9]
RP SEQUENCE OF 907-1065.
RX MEDLINE=80137544; PubMed=6987230;
RA Kingston I.B., Kingston B.L., Putnam F.W.;
RT "Primary structure of a histidine-rich proteolytic fragment of human
RT ceruloplasmin. II. Amino acid sequence of the tryptic peptides.",
RL J. Biol. Chem. 255:2886-2896(1980).
RN [10]
RP SEQUENCE OF 1007-1061 FROM N.A.
RX MEDLINE=90285218; PubMed=2355023;
RA Yang F.M., Friedrichs W.E., Cupples R.L., Banifacio M.J.,
RA Sanford J.A., Horton W.A., Bowman B.H.;
RT "Human ceruloplasmin. Tissue-specific expression of transcripts
RT produced by alternative splicing.",
RL J. Biol. Chem. 265:10780-10785(1990).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS).
RA Zaitseva I., Zaitsev V., Card G., Moshkov K., Bax B., Ralph A.,
RA Lindley P.;
RT "The X-ray structure of human serum ceruloplasmin at 3.1 A: nature of
RT the copper centres.",
RL J. Biol. Inorg. Chem. 1:15-23(1996).

```



| FT                    | CARBONHD          | 138  | 138                 | N-LINKED (GLCNAC. . .)          |
|-----------------------|-------------------|--|---------------------|---------------------------------|
| FT                    | CARBONHD          | 358  | 358                 | N-LINKED (GLCNAC. . .)          |
| FT                    | CARBONHD          | 397  | 397                 | N-LINKED (GLCNAC. . .)          |
| FT                    | CARBONHD          | 762  | 762                 | N-LINKED (GLCNAC. . .)          |
| FT                    | DISULFID          | 174  | 200                 | PROBABLE.                       |
| FT                    | DISULFID          | 276  | 357                 | PROBABLE.                       |
| FT                    | DISULFID          | 534  | 560                 | PROBABLE.                       |
| FT                    | DISULFID          | 637  | 718                 | PROBABLE.                       |
| FT                    | DISULFID          | 874  | 900                 | PROBABLE.                       |
| FT                    | METAL             | 120  | 120                 | COPPER (TYPE 2) (BY SIMILARITY) |
| FT                    | METAL             | 122  | 122                 | COPPER (TYPE 3) (BY SIMILARITY) |
| FT                    | METAL             | 180  | 180                 | COPPER (TYPE 3) (BY SIMILARITY) |
| FT                    | METAL             | 182  | 182                 | COPPER (TYPE 3) (BY SIMILARITY) |
| FT                    | METAL             | 994  | 994                 | COPPER (TYPE 1) (BY SIMILARITY) |
| FT                    | METAL             | 997  | 997                 | COPPER (TYPE 2) (BY SIMILARITY) |
| FT                    | METAL             | 999  | 999                 | COPPER (TYPE 3) (BY SIMILARITY) |
| FT                    | METAL             | 1039   | 1039                | COPPER (TYPE 3) (BY SIMILARITY) |
| FT                    | METAL             | 1040   | 1040                | COPPER (TYPE 1) (BY SIMILARITY) |
| FT                    | METAL             | 1041   | 1041                | COPPER (TYPE 3) (BY SIMILARITY) |
| FT                    | METAL             | 1045   | 1045                | COPPER (TYPE 1) (BY SIMILARITY) |
| FT                    | METAL             | 1050   | 1050                | COPPER (TYPE 1) (BY SIMILARITY) |
| FT                    | VARIANT           | 79   | 79                  | T -> G.                         |
| FT                    | VARIANT           | 449  | 449                 | /FTID=VAR_001043.               |
| FT                    | VARIANT           | 449  | 449                 | L -> G.                         |
| FT                    | CONFLICT          | 1060   | 1060                | /FTID=VAR_001044.               |
| FT                    | CONFLICT          | 1065   | 1065                | E -> EGEVP (IN REF. 4).         |
| SO                    | SEQUENCE          | 1065   | 122205              | MM; 2F2F1294E2D30F58 CRC64;     |
| Query Match           |                   | 21.9%;   | Score 1725;         | DB 1;                           |
| Best Local Similarity |                   | 32.1%;   | Pred. No. 3.8e-105; |                                 |
| Matches 388;          | Conservative 196; | Mismatches 417;  | Indels 208;         | Gaps 23                         |
| QY                    | 5                 | LSTCFCLLRPCFSATRRYVLAVALSDYQSDLGE---                           | LPVDARFPVPKSPFNT    | 61                              |
| DB                    | 6                 | LGIFLFLCSTP-AMAKKEKHYYIGIETTWY-ASDHGEKKLISVTEHSHNITYQNPDR1     |                     | 63                              |
| QY                    | 62                | SVVYKKTLEVEETVALVNIAPRPMWGLGPTIOAEVYDVTIVITLKNASHPVLSHAVG      |                     | 121                             |
| DB                    | 64                | GRVKKALYLYQDTDETFTTEKPYWLGFLGPIIAKETGDKVYVHDKNLASRYPETHSHG     |                     | 123                             |
| QY                    | 122               | VSYKASGAEEDDOTSQREKEDKVPFGGSHITTYQVYLKENGPMASDPLCLTYSLSHV      |                     | 181                             |
| DB                    | 124               | ITYYKEHGAITYPDMTTFORADKDVYRGEQYTYMLATEBOSPEEGDCNCVTRIYHSH1     |                     | 183                             |
| QY                    | 182               | DLVVDNLGGLGALLVCREGSLAKKCTQL-HKFLTFAPVPEKGSWSE-----TKNS        |                     | 234                             |
| DB                    | 184               | DAPKDIAAGLLGRLTLCKKSLDKKEKHIDREFVVMSSVDENFSWYLEDNIKTYCSEP      |                     | 243                             |
| QY                    | 235               | LMORDAASAPAMPAMHTVNGVYNSRLGGLGCHKRSYVWHYIGMTPEVHSIFLEGHT       |                     | 294                             |
| DB                    | 244               | EKVQKDNDDPESBNMYSVNGCTPESLGLGSCACADRYKMYLFGMGNEVDYNAAPFHGGA    |                     | 303                             |
| QY                    | 295               | FLVNRHROASLEISDITFLTAOTLIMDLGQFLFLFCHISHSHOHGMEAYVYKVDSCPEEPOL |                     | 354                             |
| DB                    | 304               | LTNKKYRIDTINLPATFLDAVMVAQNGEWSLSCQNLNHLKAGLDAEFQVQEC-----      |                     | 357                             |
| QY                    | 355               | RMKNNKEAEVDYDDDLTDBEMDVVRRDDNSPFIQIRSVAKKHPTWVHYITAAEEDWDY     |                     | 414                             |
| DB                    | 358               | ---NKSSSKD-----NIRKNVHN-----YYIAAEILIMY                        |                     | 385                             |
| QY                    | 415               | APL-----VLAPDDRSYKSOYLANNGPORGIRGKYKVRPMATYDETET--KTREAIO      |                     | 462                             |
| DB                    | 386               | APSGIDITFKENLAPRSGDS--AVEFEGCTTRIGSYKKLYVRETTDASFTKRKERGPEE    |                     | 443                             |
| QY                    | 463               | HESGILGPLXGEGDTLLIFKQDASRPYNIYPHGI-----TDVRYPL--SRRLP          |                     | 511                             |
| DB                    | 444               | EHLTILGPIVIAAEVGGDITRVTFFHNGAFLSIEPIGCVFRFNKNNECTYYSPNYNPQSRVP |                     | 503                             |
| QY                    | 512               | KGVNHLKDFLLPGEIEFKYKTYVDEQPTSDPRLCLTRYSSPYNMERDLASGLIGPL       |                     | 571                             |
| DB                    | 504               | PSASH-----VAPETETVYEWTVREKVGPTNADPCLAKMYYSADVDPKIDITGLIGPMK    |                     | 558                             |

| ID | CERU  | RAT                     | STANDARD:                 | PRT:           | 1059 | AA. |
|----|---|-------------------------|---------------------------|----------------|------|-----|
| AC | P13635;   | 064719;                 |                           |                |      |     |
| DT | 01-JAN-1990   | (Rel. 13,               | Created)                  |                |      |     |
| DT | 01-NOV-1997   | (Rel. 35,               | Last sequence update)     |                |      |     |
| DT | 15-JUL-1999   | (Rel. 38,               | Last annotation update)   |                |      |     |
| DE | CERULOPLASMIN   | PRECURSOR               | (EC 1.16.3.1)             | (FERROXIDASE). |      |     |
| GN | CP.   |                         |                           |                |      |     |
| OS | Rattus norvegicus   | (Rat).                  |                           |                |      |     |
| OC | Eukaryota; Metacoa; Chordata;                                       | Cranialta; Vertebrata;  | Euteleostomi;             |                |      |     |
| OC | Mammalia; Eutheria;   | Rodentia; Sciurognathi; | Muridae; Murinae; Rattus. |                |      |     |
| OX | NCBI_TaxId=10116;   |                         |                           |                |      |     |
| RN | [1]   |                         |                           |                |      |     |
| RP | SEQUENCE FROM N.A.  |                         |                           |                |      |     |
| RC | STRAIN=SPRAGUE-DAWLEY;  | TISSUE=Lung,            | and Liver;                |                |      |     |
| RA | MEDLINE=90237081;   | PubMed=2332446;         |                           |                |      |     |
| RX | Fleming R.E., Giltlin J.D.;   |                         |                           |                |      |     |
| RT | "Primary structure of rat ceruloplasmin and analysis of             |                         |                           |                |      |     |
| RT | tissue-specific gene expression during development."                |                         |                           |                |      |     |
| RL | J. Biol. Chem. 265:7701-7707(1990).                                 |                         |                           |                |      |     |
| RN | [2]   |                         |                           |                |      |     |
| RP | SEQUENCE OF 257-294;  | 571-612 AND 823-892     | FROM N.A.                 |                |      |     |
| RC | TISSUE=Liver;   |                         |                           |                |      |     |
| RA | MEDLINE=87137545;   | PubMed=3818625;         |                           |                |      |     |
| RT | Aldred A.R., Grimes A., Schretler G., Mercer J.F.B.;                |                         |                           |                |      |     |
| RT | "Rat ceruloplasmin. Molecular cloning and gene expression in liver, |                         |                           |                |      |     |
| RT | choroid plexus, yolk sac, placenta, and testis."                    |                         |                           |                |      |     |

|    |                 |   |
|----|-----------------|---|
| RE | J. Biol. Chem.  | 262:2875-2878(1987).  |
| CC | -I-             | FUNCTION: CERULOPLASMIN IS A BLUE, COPPER-BINDING (6-7 ATOMS PER MOLECULE) GLYCOPROTEIN FOUND IN PLASMA. FOUR POSSIBLE FUNCTIONS ARE FERROXIDASE ACTIVITY, AMINE OXIDASE ACTIVITY, COPPER TRANSPORT AND HOMEOSTASIS, AND SUPEROXIDE DISMUTASE ACTIVITY.   |
| CC | -I-             | FUNCTION: MAY ALSO PLAY A ROLE IN FETAL LUNG DEVELOPMENT OR PULMONARY ANTIOXIDANT DEFENSE.  |
| CC | -I-             | CATALYTIC ACTIVITY: 4 Fe(2+) + 4 H(++) + O(2) = 4 Fe(3+) + 2 H(2)O.   |
| CC | -I-             | COPACITOR: BINDS 6 CU-IONS PER MOLECULE. THIS PROEIN BELONGS TO THE MULTICOPPER OXIDASES WHICH CONTAIN THREE DISTINCT CU CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE 3 OR COUPLED BINUCLEAR.  |
| CC | -I-             | TISSUE SPECIFICITY: SYNTHESIZED IN LIVER AND SECRETED INTO THE PLASMA. ALSO CHOROID PLEXUS, YOLK SAC, PLACENTA, AND TESTIS; NOT IN STOMACH AND SMALL INTESTINE. FETAL LUNG AND LIVER.   |
| CC | -I-             | INDUCTION: BY INFLAMMATION.   |
| CC | -I-             | SIMILARITY: CONTAINS 3 F5/8 TYPE A DOMAINS; EACH IS COMPOSED OF 2 PLASTOCYANIN-LIKE REPEATS.  |
| CC |                 |   |
| CC |                 | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch">http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch</a> ). |
| DR | EMBL:           | L33869; AAA40917.1; --  |
| DR | EMBL:           | M80529; AAH5820.1; --   |
| DR | EMBL:           | J02670; AAA40814.1; ALT_SEQ.  |
| DR | EMBL:           | M14102; AAA40915.1; --  |
| DR | PIR:            | A35210; A35210.   |
| DR | PIR:            | A29564; A29564.   |
| DR | HSSP:           | P00450; IKCM.   |
| DR | InterPro:       | IIPR001117; --  |
| DR | InterPro:       | IIPR002355; --  |
| DR | Pfam:           | PF00394; Cu-oxidase; 3.   |
| DR | PROSITE:        | PS00079; MULTICOPPER_OXIDASE1; 3.   |
| DR | PROSITE:        | PS00080; MULTICOPPER_OXIDASE2; 1.   |
| KW | Oxidoreductase; | Copper; Metal-binding; Glycoprotein; Plasma; Repeat;  |
| KW | Signal.         |   |
| FT | SIGNAL          | 1 19 PROBABLE.  |
| FT | CHAIN           | 20 1059 CERULOPLASMIN.  |
| FT | DOMAIN          | 20 356 F5/8 TYPE A 1.   |
| FT | DOMAIN          | 20 199 PLASTOCYANIN-LIKE 1.   |
| FT | DOMAIN          | 208 354 PLASTOCYANIN-LIKE 2.  |
| FT | DOMAIN          | 369 712 F5/8 TYPE A 2.  |
| FT | DOMAIN          | 369 554 PLASTOCYANIN-LIKE 3.  |
| FT | DOMAIN          | 564 710 PLASTOCYANIN-LIKE 4.  |
| FT | DOMAIN          | 724 1055 F5/8 TYPE A 3.   |
| FT | DOMAIN          | 724 894 PLASTOCYANIN-LIKE 5.  |
| FT | DOMAIN          | 902 1051 PLASTOCYANIN-LIKE 6.   |
| FT | DISULFID        | 173 199 BY SIMILARITY.  |
| FT | DISULFID        | 275 356 BY SIMILARITY.  |
| FT | DISULFID        | 528 554 BY SIMILARITY.  |
| FT | DISULFID        | 631 712 BY SIMILARITY.  |
| FT | DISULFID        | 868 894 BY SIMILARITY.  |
| FT | METAL           | 120 120 COPPER (TYPE 2) (BY SIMILARITY).  |
| FT | METAL           | 122 122 COPPER (TYPE 3) (BY SIMILARITY).  |
| FT | METAL           | 179 179 COPPER (TYPE 3) (BY SIMILARITY).  |
| FT | METAL           | 181 181 COPPER (TYPE 3) (BY SIMILARITY).  |
| FT | METAL           | 988 988 COPPER (TYPE 1) (BY SIMILARITY).  |
| FT | METAL           | 991 991 COPPER (TYPE 2) (BY SIMILARITY).  |
| FT | METAL           | 993 993 COPPER (TYPE 3) (BY SIMILARITY).  |
| FT | METAL           | 1033 1033 COPPER (TYPE 3) (BY SIMILARITY).  |
| FT | METAL           | 1034 1034 COPPER (TYPE 1) (BY SIMILARITY).  |
| FT | METAL           | 1035 1035 COPPER (TYPE 3) (BY SIMILARITY).  |
| FT | METAL           | 1039 1039 COPPER (TYPE 1) (BY SIMILARITY).  |
| FT | METAL           | 1044 1044 COPPER (TYPE 1) (BY SIMILARITY).  |
| FT | CARBOHYD        | 138 138 N-LINKED (GLCNAC. . . ) (POTENTIAL).  |
| FT | CARBOHYD        | 226 226 N-LINKED (GLCNAC. . . ) (POTENTIAL).  |
| FT | CARBOHYD        | 396 396 N-LINKED (GLCNAC. . . ) (POTENTIAL).  |
| FT | CARBOHYD        | 582 582 N-LINKED (GLCNAC. . . ) (POTENTIAL).  |

FT CARBOHYD 756 756 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 920 920 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 271 271 G -> A (IN REF. 2).  
 FT CONFLICT 604 605 ED -> DN (IN REF. 2).  
 FT CONFLICT 823 823 T -> S (IN REF. 2).  
 FT CONFLICT 833 833 V -> L (IN REF. 2).  
 FT CONFLICT 868 868 C -> V (IN REF. 2).  
 FT CONFLICT 891 891 L -> R (IN REF. 2).  
 SQ SEQUENCE 1059 AA, 120840 MW, 12BA3B990A0B95E3 CRC64;

Query Match 21.3%; Score 1678; DB 1; Length 1059;  
 Best Local Similarity 32.0%; Pred. No. 4,6e-102;  
 Matches 385; Conservative 188; Mismatches 429; Indels 200; Gaps 23;

5 LSTCFCLLRCEFSATRRYIGAVELSDWQ-SDLEL-PVDARPPRPVPSKSPFNTS 62  
 6 LSALEFL-HSSLWTRREHYIGITEAVWDYASGSEKELISVDPEQSNFYLRNPPRIG 64  
 63 VYKKTLEVEFTVHLFNIAKPRPMMGLGPTIOAEVYDTVITLKMAASHPVSLHAGV 122  
 65 RYKKALETITDGTFTTIDKRAMGLGPTVKAEGDKSVNKNASRPYTFTHAGV 124  
 123 SYWKASGAEYDDQTSOREKEDKVPFGSHTYVQVLKENGPMASDPLCLTYSLSHV 182  
 125 TYTKANEGAIYPDNTDFQADDKLFPGQOYLVLRA-NEPSPGSGDNCVTRYHSHV 183  
 183 LKVDLSGLIGALVRCGSLAKERTOTL-HKFTLLFAVFPDEGKSMHETKNSLM----- 236  
 184 APKDIASGLIPCLCKSLAKEENIDEFVLMFSVDENLSMLEDNIKPFCSPE 243  
 237 -QDRDASARAPMKHTNGVYVNSRLPGLICHRKSVYMHYIGMTGTEVHSIFLEGHTF 295  
 244 KYDKNEDFOESNRMYSTINGTFGSLPGLSMCAEDRYVWYLFQMGNEVDVSELPFGAL 303  
 296 LVNRHQASLEISPTFLAOTLMDIGOFFLCHISSHODGMAVYKVDSCPEPOLR 355  
 304 TSKNHTDIILFPATLIDVSVQNPVMMSCONLNLKAGLQAFQVRDC----- 356  
 356 MKNNEADYDDDLTDSMDVYRFPDDNSPSFIOIRSAKKHPKRWYVYIAEEDWYA 415  
 357 --NRPSP--DDDIODRHY-----RR-----YVIAAEFTIWIYA 385  
 416 P-----LVLPDRRSYKSOYLNNGPORIGRKYKRYKRYMAVYDTEF---KTRBAI 461  
 386 PGGTPTFGENFTSLGSSRVFPEO---GATRIGGSYKKLRYREYIDOSTPNKRERPD 441  
 462 QHESGILPRLLYGEVDTLLIFKMQASRPYNIYPHGITDVR---PLYSRRLPKGVKHL 517  
 442 EHHGLIILPVLMAEYGDIIIVTFHNKGOPLSLIPMGVRFKENGCTYGG---PDGRSSK 498  
 518 KDFPLPEIKRYKTYVEDEGPTKSDPRCLTRYSSSFVNMERDLASGLIGPLICYRES 577  
 499 QASHAPKETTYETWYKEMGPYADVCLSKMYSGVDTLKDFTGLILPMKCKGGS 558  
 578 VDORQNMDSKRWYLFVSDENRSMYLTENIORFLPNPAGVOLEDEPOASNMHSIN 637  
 559 LLAGROVDYKERYLPAFYLDENESLLDNIKFTTAPENVKREDDFOESKMHSMN 618  
 638 GYVDSIO-LSVCLHEAVYIYLSIGAOTDFLSVFFSGYTERKHKWYEDTLTLPFSGET 696  
 619 GFMYGNLGLNMCLGESIYVYLFSSAGNEADVGYIFSGNTYLSKGERDRTANLFPKSLT 678  
 697 VFMSMENGLMILGCHNSDFNRKGTALKYSSCDKNTGVDYEDSYEDISAYLSKNAI 756  
 679 LIMPEDGSDVDECLTDDHDTGKOKRYTNOQ-----KQGFEDVLY----- 722  
 757 EPRSESONSRPSTROKOFNATPVLKHKREITRTTLOSQDEIDYDVTISVEKKEDF 816  
 723 ----- 722  
 817 DIVDENOSPRSFQKTRHYFLAAVERLMDYGMSSSPHVLNRAQSGV----- 866

DB 723 -----QGETRYIAAEVEMDYSPSRDEMELHLHQEONVSNAPLKEEF 767  
 QY 867 ---PQKRVYVQEFDDSGFTQPLRGELNHLGLGPTIRAEVDENINWTFPNQASRYS 923  
 DB 768 FIGSKYKRVYREFDSTFRQVRRABEEHGLMGLPIHADVGAKVYVFNKMATRYS 827  
 QY 924 FYSSLISVEEDROGAERKKNFVK---DNERTKYMKVOHNAAPRKDEPCKAMAYFSOV 980  
 DB 828 IHA-----HGVKRSSTVAPTLGVEVTTYIWOIPENSGAGTEDSPCLPMAYISTV 877  
 QY 981 DLEKDVHSLIGPLDVC---HTNTLNPAGROVYQVEFALFETIPDETSMYTFENMERN 1037  
 DB 878 DRVKDYLSGLIGPLIVCRKSVKYNPR-----KKEPSTLFLVDEMSWYLDNDINMY 932  
 QY 1038 CRAPCNIMEDPTFKENRFPAINCYIMDTLPGLVMAODORRKYLLSMGSENHSHIF 1097  
 DB 933 PDHPEKDKNDKEEFTESKMKHAINGMKGNLOGLTMHGDENVWYVMAAMGNDIDLTWHF 992  
 QY 1098 SGHVFTVKKKEEYKALINLPGVEFVEMLPKSKGIWRVECLIGENLHAGSTLELYS 1157  
 DB 993 HGHSTQYKHHRGHSSDVEDPFGTYQTELEMPQTGTMLLCHVTDIHAGVTTYTVLP 1052  
 QY 1158 NK 1159  
 DB 1053 NO 1054

RESULT 8  
 CERU\_MOUSE STANDARD; PRT: 1062 AA.  
 ID CERU\_MOUSE  
 AC 061147;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE CERULOPLASMIN PRECURSOR (EC 1.16.3.1) (FERROXIDASE).  
 GN CP.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Klomp L.W.J., Farhangrazi Z.S., Choi D.W., Gitlin J.D.;  
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP TISSUE SPECIFICITY.  
 RX MEDLINE=96294736; PubMed=8690795;  
 RA Klomp L.W.J., Farhangrazi Z.S., Dugan L.L., Gitlin J.D.;  
 RL "Ceruloplasmin gene expression in the murine central nervous system.";  
 RL J. Clin. Invest. 98:207-215(1996).  
 CC -!- FUNCTION: CERULOPLASMIN IS A BLUE, COPPER-BINDING (6-7 ATOMS PER MOLECULE) GLYCOPROTEIN FOUND IN PLASMA. FOUR POSSIBLE FUNCTIONS ARE FERROXIDASE ACTIVITY, AMINE OXIDASE ACTIVITY, COPPER TRANSPORT AND HOMEOSTASIS, AND SUPEROXIDE DISMUTASE ACTIVITY.  
 CC -!- CATALYTIC ACTIVITY: 4 FE(2+) + 4 H(+) + O(2) = 4 FE(3+) + 2 H(2)O.  
 CC -!- COFACTOR: BINDS 6 CU-IONS PER MOLECULE. THIS PROTEIN BELONGS TO THE MULTICOPPER OXIDASES WHICH CONTAIN THREE DISTINCT CU CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE 3 OR COUPLED BINUCLEAR.  
 CC -!- TISSUE SPECIFICITY: MANY TISSUES, INCLUDING LIVER, EYE AND BRAIN.  
 CC -!- SIMILARITY: CONTAINS 3 FS/8 TYPE A DOMAINS; EACH IS COMPOSED OF 2 PLASTOCYANIN-LIKE REPEATS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U049430; AAB07996.1; -;  
 CC HSSP; P00450; IKCW.

DR MGD; MGI:88476; Cp.  
 DR InterPro; IPR001117; -  
 DR InterPro; IPR002355; -  
 DR Pfam; PF00394; Cu-oxidase; 3.  
 DR PROSITE; PS00079; MULTICOPPER\_OXIDASE2; 1.  
 DR PROSITE; PS00080; MULTICOPPER\_OXIDASE2; 1.  
 DR Oxidoreductase; Copper; Metal-binding; glycoprotein; Plasma; Repeat;  
 KW Signal.  
 FT SIGNAL 1 19  
 FT CHAIN 20 1062  
 FT DOMAIN 20 356  
 FT DOMAIN 20 199  
 FT DOMAIN 208 356  
 FT DOMAIN 369 713  
 FT DOMAIN 369 713  
 FT DOMAIN 565 713  
 FT DOMAIN 725 1057  
 FT DOMAIN 725 896  
 FT DOMAIN 904 1057  
 FT DISULFID 173 199  
 FT DISULFID 275 356  
 FT DISULFID 529 555  
 FT DISULFID 632 713  
 FT DISULFID 870 896  
 FT METAL 120 120  
 FT METAL 122 122  
 FT METAL 179 179  
 FT METAL 181 181  
 FT METAL 980 980  
 FT METAL 993 993  
 FT METAL 995 995  
 FT METAL 1035 1035  
 FT METAL 1036 1036  
 FT METAL 1037 1037  
 FT METAL 1041 1041  
 FT METAL 1046 1046  
 FT CARBOHYD 138 138  
 FT CARBOHYD 226 226  
 FT CARBOHYD 583 583  
 FT CARBOHYD 625 625  
 FT CARBOHYD 757 757  
 FT CARBOHYD 922 922  
 SQ SEQUENCE 1062 AA; 121159 MW; F3F5ZED09A238F16 CRC64;  
 Query Match 20.8%; Score 1639; DB 1; Length 1062;  
 Best Local Similarity 31.7%; Pred. No. 1.6e-99;  
 Matches 383; Conservative 188; Mismatches 437; Indels 200; Gaps 24;  
 QY 10 FLCLRFCE-----SATRRYLAGVELSWDMQ--SDGLRPVDARPPRPVKSPFN 60  
 DB 3 FLLESTFIFLYSSLALARKHYFICITAEVMDASGTEKKLISVDTEQSNFLQNGPDR 62  
 QY 61 TSVVKKTLFVEFVYHLENIARPRPMGLIPTQAEVYDVVITLLKNASHPVSIAH 120  
 DB 63 IGRYKKAIFYERTDGTSTKIDTPAWLGLFGLVIAAEVDKYYVHLKNLASIYTFHAH 122  
 QY 121 GVSVMKASGEAYDDQTSOREKEDKVPFGSHYVQVYLKENGPMASDPLCTYSYLH 180  
 DB 123 GYTYKYRYGAYYPRNTTDFQRADKVLPGQOYVYVLAH-NEPSPEGDSNCTRIYHSH 181  
 QY 181 VDLVKDNLGSLGALLVREGSLAKERTQT-LHKFLLFAVFDEGKSWHSETKNSLM-- 236  
 DB 182 VDAPKDIASGLIGPLILCKKSLYKKEKNIDQEFVLFMSVYDENLSWYLEDNIKTFCS 241  
 QY 237 ---QORDASASARMPKMHVNGVNSLPGLGCHKRSYVWHYIGCTTPEVHSIFLEGH 293  
 DB 242 PEKVVKDEDFQESNMYSINGTTFGLSPLGSCADRYKAWYLFEGNGENEDVASAFPHQ 301  
 QY 294 TFLVNRHROASLEISITFLTAOTLMDLGFLLFCHISHSHQDGEAYVKKVDSCEEPQ 353  
 DB 302 ALTSRNYQITDITNLPATLIDAYMAQNPGVWMLSCONLHLKAGLQAFQVODC----- 356

QY 354 LRKKNNEAEDYDDDLTSEMDYVREDDONSPSFQIRSVAKKHPTWYHIAAEEEDMD 413  
 DB 357 -----NKSSKD-----NIRKQVH-----XYIAAEVIVN 383  
 QY 414 YAP-LVLAADD-----SYKSOYLANNGPQIRKTKKRFMAVYDTEF--KTRALON 463  
 DB 384 YAPSGIDIFTEEKRLTASGSDSGVFEGQARRIGSYKKAAYREYTGSTNRKERGPDE 443  
 QY 464 ESGILGLTGEAGDPTLLIFKKNQASRPYNIYPHGIT-----DVRPLYSRLPRGVYHLKD 519  
 DB 444 HLGILGLVIAEVDGTTIKYTFHNKGQHLSTQMGVSFTAEANGTYTG---PPGASSQQA 500  
 QY 520 FP-ILPGEIFKRYWYTVEDGPKSDPCLTRYYSFVNERDLASGLIGPLILCYKESV 578  
 DB 501 ASHAPKXFTYETWYVPEKMGPTIYADPVCLSKMYSAVPTKDIIFGLIGPMKKCKSL 560  
 QY 579 DQGNQIMSKRNVILFSPVDEKRSWYLPENIORFLPNRAGVOLDEPDEQASINIMHSING 638  
 DB 561 LADGRQADVDEKFFLEFVYDENESLDDNIMFTHADPDQVDEDEDFQESNKNHSMNG 620  
 QY 639 YVFDLQTL-SVCLHEVAYWYILSIGAQDPLSVFSGYTFKHKMYVEDTLTFPESGETV 697  
 DB 621 FMYGNGSWPMKCLGESIVWYLFESAGNEADVHGIFESGNTYLCGEERDTANLFPKHSFL 680  
 QY 698 FMSMENPGLMILCCHNSDFRNKMTALLKVSQDKNTGDIYEDSYDISAYLLSKNALE 757  
 DB 681 LMPDFTGTFDEVLCLTTHDYTGKMKQKYVNC----- 713  
 QY 758 PRFSQNSRNPSTRQOFNATPVYLRHQREITRTLOSDEIDYDITISVEMKEDPD 817  
 DB 714 -----ORQF----- 721  
 QY 818 IYDEENQSPRSFOKTRHYEIAAVERLMDYGMSSPHVLRNRAQSGV-----EDPT 721  
 DB 722 VYIGE-----RTYVAVEVENDYSPSRAMEKELHLIDQNSVNFLLKKEFF 769  
 QY 867 --DQKRVYQFEPTDGSFTQPL-YRGELNHLGLCPYIRAEVEDINIMYTFRNQASRPYS 923  
 DB 770 IGSRYKKVYVQRTDSSFRQVYKRAEEDENHIGLIPPIHANVGDKVYVFKMATRPYS 829  
 QY 924 FYSLSLYSEDDQAGAPRKNFY---KPNETKYFFKQVGHNAHPKDEDFCKAMAYFSY 980  
 DB 830 IHA-----HGVTESSTVPTLPGEVATYVQIERSGAGEDSACIPMAVYSTV 879  
 QY 981 DLEKDVHSLIGPLILVC---HTNLPNARGROYTVOEFALEPTIDETSQVFTENMERN 1037  
 DB 880 DRYKDLISGLIGPLIYCRKSYAVKFSK-----KKNHEFLFLVPEBNSWVLDNIRKY 934  
 QY 1038 CRAPCNIQMEDPTFKENYRPHAINGYIMDTLPGLVNAQDQRIRWILSNGSNENHSIH 1097  
 DB 935 SEHPREYVKNQNEEFLESNKKHAIENGKMFGLQGLTHNVKDEVWVYLMGKGNEDIDHTYHF 994  
 QY 1098 SGHVFYRKKEEKYKMLYNLYPGVEFVEMLPKAIMVVECLIGLHLAGSTLFIYS 1157  
 DB 995 HGHSFOYKKNRGVYSDVDFLPCTYOTLEKFPOTPGTWLHCHVYDHAAGNATYTVLP 1054  
 QY 1158 NKCOTPLG 1165  
 DB 1055 VEGETKSG 1062  
 RESULT 9  
 MFGM\_RAT  
 ID MFGM\_RAT STANDARD; PRT: 427 AA.  
 AC P70490;  
 DT 01-NOV-1997 (rel. 35, Created)  
 DT 01-NOV-1997 (rel. 35, Last sequence update)  
 DT 30-MAY-2000 (rel. 39, Last annotation update)  
 DE LACTADHERIN PRECURSOR (MILK FAT GLOBULE-BGF FACTOR 8) (MFG-88) (O-  
 ACETYL GD3 GANGLIOSIDE SYNTHASE) (AGS) (MFGM).  
 GN MFG8 OR AGS.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Brain;
RX MEDLINE=96374422; PubMed=8780713;
RA Ogura K., Nara K., Watanabe Y., Kohno K., Tai T., Sanai Y.;
RT "Cloning and expression of cDNA for O-acetylation of GD3
   ganglioside";
RL Biochem. Biophys. Res. Commun. 225:932-938(1996).
CC -I- FUNCTION: MAY BE INVOLVED IN PHOSPHOLIPID BINDING. SEEMS TO
CC PARTICIPATE IN THE O-ACETYLATION OF GD3 GANGLIOSIDE SIALIC ACID.
CC -I- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN.
CC -I- TISSUE SPECIFICITY: SPLEEN, LUNG, HEART, BRAIN AND MUSCLE.
CC -I- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -I- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: D84068; BAA12210.1; -.
DR HSSP: P00740; IIXA.
DR InterPro: IPR000421; -.
DR InterPro: IPR000561; -.
DR InterPro: IPR001438; -.
DR Pfam: PF00008; EGF_2.
DR Pfam: PF00754; F5_F8_type_C_2.
DR PRINTS: PR00010; EGFBL00D.
DR PROSITE: PS00022; EGF_1; 2.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01285; FA58C_1; 2.
DR PROSITE: PS01286; FA58C_2; 2.
DR PROSITE: PS01286; FA58C_2; 2.
KW Signal. Glycoprotein; Repeat; EGF-like domain; Milk.
FT SIGNAL 1 22
FT CHAIN 23 427
FT DOMAIN 24 61
FT DOMAIN 64 108
FT DOMAIN 111 267
FT DOMAIN 272 427
FT DOMAIN 28 39
FT DISULFID 33 49
FT DISULFID 51 60
FT DISULFID 68 79
FT DISULFID 73 96
FT DISULFID 98 107
FT DISULFID 111 267
FT DISULFID 254 258
FT DISULFID 272 427
FT SITE 87 89
FT CARBOHYD 61 61
FT CARBOHYD 230 230
FT CARBOHYD 280 280
FT CARBOHYD 390 390
FT CARBOHYD 427 427
FT SEQUENCE 427 AA; 47413 MW; EA8C8631F3EB047 CRC64;
-----
Query Match 8.4%; Score 663; DB 1: Length 427;
Best Local Similarity 40.3%; Pred. NO. 3.7e-36;
Matches 142; Conservative 61; Mismatches 123; Indels 26; Gaps 6;
-----
OY 1138 ECLIEHLAGMSTFLVYSNK-----CPTPLGASGHIRDFQTASGQY-- 1182
DB 78 KCLVTEDEPQRG--DIFTEYICQCPVGYGSHCELGCKTKLGEGLADSDQSASASYMG 135
OY 1183 ----GOMAPKLARLHYSGSINAW--STKEPFWIKVLDLAPMIHGKTKGARQKRFSSLY 1236
DB 136 FMGLQRMGPFLARLYRTGIVNATWASSYDSDPKMIGVDFLRKKRWAGVMTQGSARAGRAEY 195

```

```

OY 1237 ISOFTIMYSLDKKQWQYTKGNSTGLMWFPGVNDSSGINKHFNPNPIARYIRLPHTHYS 1296
DB 196 LKTFKVAVSLDGRREFEIDESGTGDKHEFMGNODNNSLKINNFNPLEQYIRLVPSC 255
OY 1297 IRTSLRMLMGCDLNSCSNPLCOMESKAISDAQITASSYFT--NMFA-TWSPSKARLHIOG 1353
DB 256 RCTLRFLFELLCGELGHCSEPLGLKNNITPDSSQITFASSSYKWNLLRAFGVYPIHGRIDNG 315
OY 1354 RSNARPOVNNKEMLYOVDFQKTMVGTGTQGVASLSLTMVYKFEILSSODGHOPTLF 1413
DB 316 KINAMTAOSNAKEMLYOVLDTGOKRYVTGITTGARDPFGHIQYVASTKYAHSDDGVQWTVY 375
OY 1414 FONGKVKVFOGQDSFTPVVNSLDPLLRIRYRIRHPOSWVHOIALRMEVLC 1465
DB 376 EEOGTSKVPQSLDNNSHKKNIFKPFMARVRYVPLSMHNRITRLLELGC 427
-----
RESULT 10
MFGM_MOUSE STANDARD: PRT: 463 AA.
AC P21956; P87800;
DT 01-AUG-1991 (rel. 19, Created)
DT 01-AUG-1991 (rel. 19, Last sequence update)
DT 01-OCT-2000 (rel. 40, Last annotation update)
DE LACTADHERIN PRECURSOR (MILK FAT GLOBULE-EGF FACTOR 8) (MFG-E8) (MFGM)
DE (SPEARM SURFACE PROTEIN SP47) (MP47).
GN MFGEB8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-35.
RC TISSUE-Mammary gland;
RX MEDLINE=91046008; PubMed=2122462;
RA Stubbs J.D., Lekutis C., Singer K.L., Bul A., Yuzuki D.,
RA Srinivasan U., Parry G.;
RT "cDNA cloning of a mouse mammary epithelial cell surface protein
RT reveals the existence of epidermal growth factor-like domains linked
RT to factor VIII-like sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 87:8417-8421(1990).
RL [2]
RP SEQUENCE OF 23-463 FROM N.A.
RC TISSUE-Testis;
RA Enselin M.A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: MAY BE INVOLVED IN PHOSPHOLIPID BINDING. ZONA PELLUCIDA-
CC BINDING PROTEIN.
CC -I- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN.
CC -I- TISSUE SPECIFICITY: MAMMARY EPITHELIAL CELL SURFACES AND
CC SPERMATOZOAN.
CC -I- DEVELOPMENTAL STAGE: MRNA EXPRESSION IS DETECTABLE IN MAMMARY
CC TISSUE FROM NONPREGNANT ANIMALS & MAXIMAL IN THE LACTATING GLAND.
CC -I- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -I- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: M38337; AAA39534.1; -.
DR EMBL: Y11684; CAA72380.1; -.
DR PIR: A36479; A36479.
DR HSSP: P00740; IIXA.
DR MGD: MGI:102768; Mige8.
DR InterPro: IPR000421; -.
DR InterPro: IPR000561; -.
DR InterPro: IPR001438; -.
DR Pfam: PF00008; EGF_2.

```

DR Pfam: PF00754; F5\_F8\_type\_C; 2.  
DR PRINTS: PR00010; EGFLOOD.  
DR PROSITE: PS00022; EGF\_1; 2.  
DR PROSITE: PS01186; EGF\_2; 2.  
DR PROSITE: PS01285; FA58C\_1; 2.  
DR PROSITE: PS01286; FA58C\_2; 2.  
KW Signal; Glycoprotein; Repeat; EGF-like domain; Milk.  
FT SIGNAL 1 22  
FT CHAIN 23 463 LACTADHERIN.  
FT DOMAIN 24 61 EGF-Like 1.  
FT DOMAIN 64 108 EGF-Like 2.  
FT DOMAIN 148 303 F5/8 Type C 1.  
FT DOMAIN 308 463 F5/8 Type C 2.  
FT SITE 87 89 CELL ATTACHMENT SITE (POTENTIAL).  
FT DISULFID 28 39 BY SIMILARITY.  
FT DISULFID 33 49 BY SIMILARITY.  
FT DISULFID 51 60 BY SIMILARITY.  
FT DISULFID 68 79 BY SIMILARITY.  
FT DISULFID 73 96 BY SIMILARITY.  
FT DISULFID 98 107 BY SIMILARITY.  
FT DISULFID 148 303 BY SIMILARITY.  
FT DISULFID 308 294 BY SIMILARITY.  
FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 266 266 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 316 316 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 426 426 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CONFLICT 30 30 S -> F (IN REF. 2).  
FT CONFLICT 35 35 S -> D (IN REF. 1; AA SEQUENCE).  
FT CONFLICT 110 147 EFNYYNLDGEVFTAVPTAVPTAPTPAFTDLSNNLASR ->  
G (IN REF. 2).  
FT CONFLICT 168 168 Y -> S (IN REF. 2).  
FT CONFLICT 196 196 L -> T (IN REF. 2).  
FT CONFLICT 309 309 L -> S (IN REF. 2).  
FT CONFLICT 395 395 E -> A (IN REF. 2).  
SQ SEQUENCE 463 AA; 51465 MW; D7B8C6EFPB8A724D CRC64;  
Query Match 8.3%; Score 657; DB 1; Length 463;  
Best Local Similarity 42.2%; Pred. No. 1e-35;  
Matches 135; Conservative 63; Mismatches 110; Indels 12; Gaps 5;  
QY 1157 SNKQTPLGMSCHIRDPQITASGQY-----GQMAPKLARLHYSISINAM--STKEPS 1208  
DB 145 ASRCSTQJMGEGALADSGISASYVMGFMGLQRMGPRLARLYRTSIVAMAHNSYDSKP 204  
QY 1209 WIKVDLAPMIITIKTQGARQKFSLYISOFIIMYSLDGKMQYRGNSTGLMVFQGN 1268  
DB 205 WIOVNLKMRVSGVMTQGSARAGRAEYLKTFKVASLDGRKFEFIQDESGD-KEFLGN 263  
QY 1269 VDSGKINNIENPPIIARYIRLPHYSIRSLRMELMGCDLNSCMPLGMEKSAISDAQ 1328  
DB 264 LDNNSLTKNMFPTLEAGYIRLYPVSCHRGCTLRFELLCELLHGCLEPLGKLNNTIPDSQ 323  
QY 1329 ITASSYFT--NMFA-TWSSKRLHLQGRSNMARPQVNPKEQLQVDFOKTAKVTVGTQ 1385  
DB 324 MSASSSYTWNIRAFGMVPHGLRLNQKINAMTAQNSAKMQLQVLDLCTQRYVTGILITQ 383  
QY 1386 GVKSLTSMYVEFLISSQDGHQWTLFQNGKVKVFGQGNQDSFTPVVNSLDPLLTLYL 1445  
DB 384 GARDGCHQYVESKYVAISDDQVMTVVEQSSKVFQGNLDNNSHKHNIIEKPFMARVY 443  
QY 1446 RIHPQSWYHQIALRMEVLGC 1465  
DB 444 RVLPSVSHNRITRLRLLELGC 463  
RESULT 11  
MEGM\_PIG STANDARD: PRT; 409 AA.  
AC P79385;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE LACTADHERIN (MILK FAT GLOBULE-EGF FACTOR 8) (MEG-E8) (MEGM) (SPERM  
GN SURFACE PROTEIN SP47) (PP47).  
GN MEG8.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC TISSUE-Tests;  
RA Ensslin M.A.;;  
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: MAY BE INVOLVED IN PHOSPHOLIPID BINDING. ZONA PELLUCIDA-  
CC BINDING PROTEIN.  
CC -1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN.  
CC -1- TISSUE SPECIFICITY: MAMMARY EPITHELIAL CELL SURFACES AND  
CC SPERMATOZOAN.  
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.  
CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
CC or send an email to [license@isb.sib.ch](mailto:license@isb.sib.ch)).  
CC -----  
DR EMBL: Y11683; CAA72379.1; -  
DR HSSP: P00740; IIXA.  
DR InterPro: IPR000421; -  
DR InterPro: IPR000561; -  
DR Pfam: PF00008; EGF\_2.  
DR Pfam: PF00754; F5\_F8\_type\_C; 2.  
DR PROSITE: PS00022; EGF\_1; 2.  
DR PROSITE: PS01186; EGF\_2; 2.  
DR PROSITE: PS01285; FA58C\_1; 2.  
DR PROSITE: PS01286; FA58C\_2; 2.  
KW Glycoprotein; Repeat; EGF-like domain.  
FT DOMAIN 2 41 EGF-Like 1.  
FT DOMAIN 44 88 EGF-Like 2.  
FT DOMAIN 91 247 F5/8 Type C.  
FT DOMAIN 252 409 F5/8 Type C.  
FT SITE 67 69 CELL ATTACHMENT SITE (POTENTIAL).  
FT DISULFID 6 17 BY SIMILARITY.  
FT DISULFID 11 29 BY SIMILARITY.  
FT DISULFID 31 40 BY SIMILARITY.  
FT DISULFID 91 247 BY SIMILARITY.  
FT DISULFID 234 238 BY SIMILARITY.  
FT DISULFID 252 409 BY SIMILARITY.  
FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 409 AA; 45725 MW; BOC07AF80029927A CRC64;  
Query Match 8.3%; Score 650; DB 1; Length 409;  
Best Local Similarity 39.3%; Pred. No. 2.5e-35;  
Matches 139; Conservative 62; Mismatches 125; Indels 28; Gaps 6;  
QY 1138 ECLIGELHLAGSTFLVYSNK-----GQPLGMSCHIRDPQITASGQY-- 1182  
DB 58 ECEVIDDAHRG--DVETETICCKPHGYTGICCEIILNAPLGMEGAIDAFQISASSMHLG 115  
QY 1183 ----GQMAPKLARLHYSISINAM--STKEPSWIKVDLAPMIITIKTQGARQKFS 1236  
DB 116 FMGLQRMAPELRLHRAGLVNAMTASVNDRAVPIQVNLRRKRVGVVYQGSARGSAEY 175  
QY 1237 ISOFIIMYSLDGKMQYRGNSTGLMVFQGNVDSGKIKHNIPPIIARYIRLPHYS 1296  
DB 176 MKTFKVAYSTDRKRFQFIQGAESGDKIFMGNLNSGLKYNLFEPVLELVQYVRLPIICH 235  
QY 1297 INSTLRMELMGCDLNSCMPLGMEKSAISDAQITASSYFTN--MFATWSPSKARLHLQ 1353

|            |  |  |              |
|------------|--|--|--------------|
| Db         | 236  | RCGLTRFRELLEGCSELGSAEPLGLKMDTINPKQITASSFRTMGSLAFSMYPYVALDNG    | 295          |
| Qy         | 1354   | RSNAREQVNNPKEMLVQDVFQKTMKVTGYTQGVSKLTSMYVKEFLSSODGHQWTLF       | 141313       |
| Db         | 296  | KFNAMTQSNASASBWLQIDLGOSORVATGILITOGARDFGHQIQVAAVKAAYSDDGSVMTEY | 355          |
| Qy         | 1414   | FQNGKV--KVFQGNODSFPPVVASLDPELTLRRLRHPSQWVHQLAKREVLGC           | 1465         |
| Db         | 356  | RDQALGKLTFFGNLDNHSKKMFFETPLTLFREVRLIPYAMHNRIITLRLVELLGC        | 409          |
| RESULT 12  |  |  |              |
| MFGM_BOVIN |  |  |              |
| ID         | MFGM_BOVIN   | STANDARD:  | PRT: 427 AA. |
| AC         | Q95114; Q27959; P79344;  |  |              |
| DT         | 01-NOV-1997 (Rel. 35, Created)   |  |              |
| DT         | 01-NOV-1997 (Rel. 35, last sequence update)                            |  |              |
| DT         | 01-OCT-2000 (Rel. 40, last annotation update)                          |  |              |
| DE         | LACTADHEIN PRECURSOR (MILK FAT GLOBULE-EGF FACTOR 8) (MFG-E8)          |  |              |
| DE         | (MGF57/53) (PAS-6/PAS-7 GLYCOPROTEIN) (MFGM) (SPERM SURFACE PROTEIN    |  |              |
| GN         | SE47) (BP47) (COMPONENTS 15/16).                                       |  |              |
| GN         | MFG8.  |  |              |
| OS         | Bos taurus (Bovine).   |  |              |
| OC         | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;      |  |              |
| OC         | Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;      |  |              |
| OC         | Bovidae; Bovineae; Bos.  |  |              |
| OX         | NCBI_TaxID=9913;   |  |              |
| RN         | [1]  |  |              |
| RP         | SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CARBOHYDRATE-LINKAGE SITES.  |  |              |
| RC         | SPRAIN=HOLSTEIN; TISSUE=Mammary gland.                                 |  |              |
| RX         | MEDLINE=97008954; PubMed=8836004;                                      |  |              |
| RA         | Hvaerregaard J., Andersen M.H., Berglund L., Rasmussen J.T.,           |  |              |
| RA         | Petersen T.E.;   |  |              |
| RT         | "characterization of glycoprotein PAS-6/7 from membranes of bovine     |  |              |
| RT         | milk fat globules."  |  |              |
| RL         | Eur. J. Biochem. 240:628-636(1996).                                    |  |              |
| RN         | [2]  |  |              |
| RP         | SEQUENCE OF 18-427 FROM N.A.   |  |              |
| RC         | TISSUE=Mammary gland;  |  |              |
| RX         | MEDLINE=96125736; PubMed=8541316;                                      |  |              |
| RA         | Aoki N., Kishi M., Taniguchi Y., Adachi T., Nakamura R.,               |  |              |
| RA         | Matsuda T.;  |  |              |
| RT         | "Molecular cloning of glycoprotein antigens MGF57/53 recognized by     |  |              |
| RT         | monoclonal antibodies raised against bovine milk fat globule           |  |              |
| RT         | membrane."   |  |              |
| RL         | Biochim. Biophys. Acta 1245:385-391(1995).                             |  |              |
| RN         | [3]  |  |              |
| RP         | SEQUENCE OF 19-427 FROM N.A.   |  |              |
| RC         | TISSUE=Testis;   |  |              |
| RA         | Ensslin M.A.;  |  |              |
| RT         | Submitted (NOV-1997) to the EMBL/genbank/DBJ databases.                |  |              |
| RN         | [4]  |  |              |
| RP         | SEQUENCE OF 140-146; 174-187; 233-246 AND 422-427.                     |  |              |
| RC         | TISSUE=Milk;   |  |              |
| RX         | MEDLINE=93250576; PubMed=8485470;                                      |  |              |
| RA         | Mather I.H., Banghart L.R., Lane W.S.;                                 |  |              |
| RT         | "The major fat-globule membrane proteins, bovine components 15/16 and  |  |              |
| RT         | guinea-pig GP 55, are homologous to MGF-E8, a murine glycoprotein      |  |              |
| RT         | containing epidermal growth factor-like and factor V/VIII-like         |  |              |
| RT         | sequences."  |  |              |
| RL         | Biochem. Mol. Biol. Int. 29:545-554(1993).                             |  |              |
| CC         | -1- FUNCTION: PROBABLY ASSOCIATES WITH PHOSPHOLIPIDS ON THE SURFACE OF |  |              |
| CC         | MAMMARY EPITHELIAL CELLS AND MILK FAT GLOBULES. ZONA PELLUCIDA-        |  |              |
| CC         | BINDING PROTEIN.   |  |              |
| CC         | -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A   |  |              |
| CC         | SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING. THE SHORT FORM       |  |              |
| CC         | LACKS 53 AMINO ACIDS WITHIN THE F5/8 TYPE C 1 DOMAIN.                  |  |              |
| CC         | -1- TISSUE SPECIFICITY: MILK AND SPERMATOZOAN.                         |  |              |
| CC         | -1- PTM: THE 2 O-LINKED GLYCANS CONSIST OF GAL, GLCNAC AND FUC, WITH   |  |              |
| CC         | PROBABLY FUC AS REDUCING TERMINAL SUGAR.                               |  |              |
| CC         | -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.                           |  |              |
| CC         | -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.                        |  |              |

[illegible]



DB 312 OCKENAMTAQTSASAEWLOIDIGSQKRVGICITOGARDGCHIOYVAAYRVAGDDGYTWT 371  
QY 1412 LEFONG--KVKYFOGNDSEFTPVNSLDPLLTRYLRIHPQSWVHOJALREMYLGC 1465  
DB 372 EKKDQGASCKIFPGMNMNHSKKNIIEPDPQARFVRIQPVAMHNRTTLVELLGC 427  
RESULT 13  
MFGM\_HUMAN STANDARD; PRT; 387 AA.  
AC 008431;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE LACTADHERIN PRECURSOR (MILK FAT GLOBULE-EGF FACTOR 8) (MFG-EB) (HMF8)  
DE (BREAST EPITHELIAL ANTIGEN BA46) (MFGM) [CONTAINS: MEDIN].  
GN MFGEB.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Breast, and Breast carcinoma;  
RX MEDLINE=96213908; PubMed=8639264;  
RA Couto J.R., Taylor M.R., Godwin S.G., Ceriani R.L., Peterson J.A.;  
RT "Cloning and sequence analysis of human breast epithelial antigen  
BA46 reveals an RGD cell adhesion sequence presented on an epidermal  
growth factor-like domain.";  
RT DNA Cell Biol. 15:281-286(1996).  
RN [2]  
RP SEQUENCE OF 170-387 FROM N.A.  
RC TISSUE-Mammary gland;  
RX MEDLINE=91371351; PubMed=1909932;  
RA Larocca D., Peterson J.A., Urra R., Kuniyoshi J., Bistrain A.M.,  
RT "Isolation and characterization of full and truncated forms of human  
breast carcinoma protein BA46 from human milk fat globule membranes.";  
RT J. Protein Chem. 17:143-148(1998).  
RN [3]  
RP PARTIAL SEQUENCE, AND CHARACTERIZATION.  
RC TISSUE-Milk;  
RX MEDLINE=98194924; PubMed=9535276;  
RA Gluffrida M.G., Cavaleto M., Giunta C., Conti A.,  
RT "Isolation and characterization of full and truncated forms of human  
breast carcinoma protein BA46 from human milk fat globule membranes.";  
RT J. Protein Chem. 17:143-148(1998).  
RN [4]  
RP SEQUENCE OF 268-317, AND IDENTIFICATION OF MEDIN.  
RX MEDLINE=99342076; PubMed=10411933;  
RA Haeggyist B., Naejund J., Sletten K., Westermarck G.T., Mucchiano G.,  
RT "Isolation and characterization of full and truncated forms of human  
breast carcinoma protein BA46 from human milk fat globule membranes.";  
RT J. Protein Chem. 17:143-148(1998).  
RN [5]  
RP CHARACTERIZATION.  
RX MEDLINE=97405885; PubMed=92260929;  
RA Taylor M.R., Couto J.R., Scallan C.D., Ceriani R.L., Peterson J.A.;  
RT "Lactadherin (formerly BA46), a membrane-associated glycoprotein  
expressed in human milk and breast carcinomas, promotes Arg-Gly-Asp  
(RGD)-dependent cell adhesion.";  
RT DNA Cell Biol. 15:861-869(1997).  
RN [6]  
RP FUNCTION: MAY BE INVOLVED IN PHOSPHOLIPID BINDING. BINDS  
SPECIFICALLY TO ROTAVIRUS AND INHIBITS ITS REPLICATION.  
CC -1- FUNCTION: MEDIN IS THE MAIN CONSTITUENT OF AORTIC MEDIAL AMYLOID.  
CC -1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN.  
CC -1- TISSUE SPECIFICITY: MAMMARY EPITHELIAL CELL SURFACES AND AORTIC  
MEDIA. OVEREXPRESSED IN SEVERAL CARCINOMAS.  
CC -1- PTM: MEDIN HAS A RAGGED N-TERMINUS WITH MINOR SPECIES STARTING AT  
AMINO ACID 264 AND 273.  
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
CC or send an email to [license@isb.ch](mailto:license@isb.ch)).  
CC -----  
DR EMBL: U58516; AAC50549.1; -  
DR EMBL: S56151; AAB19771.1; -  
DR MM: 602281; -  
DR InterPro: IPR000421; -  
DR InterPro: IPR000561; -  
DR Pfam: PF00008; EGF\_1.  
DR Pfam: PF00754; F5\_F8\_type\_C\_2.  
DR PROSITE: PS00022; EGF\_1; 1.  
DR PROSITE: PS01186; EGF\_2; 1.  
DR PROSITE: PS01285; FA58C\_1; 2.  
DR PROSITE: PS01286; FA58C\_2; 2.  
DR Signal: Glycoprotein; Milk; Repeat; EGF-like domain; Amyloid.  
FT SIGNAL 1 23  
FT CHAIN 24 387 LACTADHERIN, SHORT FORM.  
FT CHAIN 202 387 LACTADHERIN, SHORT FORM.  
FT CHAIN 268 317 MEDIN.  
FT DOMAIN 24 67 EGF-LIKE.  
FT DOMAIN 70 225 F5/8 TYPE C 1.  
FT DOMAIN 230 387 F5/8 TYPE C 2.  
FT SITE 46 48 CELL ATTACHMENT SITE (POTENTIAL).  
FT DISULFID 27 38 BY SIMILARITY.  
FT DISULFID 32 55 BY SIMILARITY.  
FT DISULFID 57 66 BY SIMILARITY.  
FT DISULFID 70 225 BY SIMILARITY.  
FT DISULFID 212 216 BY SIMILARITY.  
FT DISULFID 387 BY SIMILARITY.  
FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 325 325 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SO SEQUENCE 387 AA; 43123 MW; 2EE6571DEC83782D CRC64;  
Query Match 7.5%; Score 588; DB 1; Length 387;  
Best Local Similarity 37.3%; Pred. No. 2,6e-31;  
Matches 132; Conservative 69; Mismatches 125; Indels 28; Gaps 9;  
QY 1126 EMPSKAGIMRVECLGHEHLHAGMSFLVYSKQOTPLGMSGHRDQITASG----- 1180  
DB 48 DVFP-----YTCYCLKG--YAGNHC-----ETKVEPLGMEGNTANSQIAASSVRYTF 95  
QY 1181 -QYGVAPKRLARLYHSINAW--SYKEPFSWIKVDLAPMIIHGKTGARKFSSLYI 1237  
DB 96 LGIQHVPPELARIINRAGMNAWNPSSNDNPMLQVALLRRMWTGVYVYGASRLASHETL 155  
QY 1238 SQPTIYSLDGKKWQYTRGNSGTGLVYFEGVNDSSGKIKINFPRIARYIRLPHTHYSI 1297  
DB 156 KAKFVAYSLNGHGF-D-FIDVYNKKHKEFVGNNKNAVHNLETPPEADQVYLYPTSCHT 214  
QY 1298 RSTLRMLMGCDLNSCMPLGMSKAIQDQITASSYF-----TNMFATSPSKARLHLDG 1353  
DB 215 ACLRLRELLGCELNGCANPLGLKNNISIPDKQITASSYKTWGLHIF-SNPPYARLDDQG 273  
QY 1354 RSNAPVYNNPKEWLOVDFOKTKMYGTGQVSLTSMVKEFLISSDQDGHQWTLF 1413  
DB 274 NFNAMWAGSYGNDQWLOVDFGSSKEVYGLITOGARFSGVDFVASTKAYISDASNMYEY 333  
QY 1414 F--QNGKVVVFOGNDSEFTPVNSLDPLLTRYLRIHPQSWVHOJALREMYLGC 1465  
DB 334 QDRPTGSSIFPGMNMNHSKKNIIEPDPQARFVRIQPVAMHNRTTLVELLGC 387  
RESULT 14

[illegible]

RESULT 15  
 NRP2\_RAT STANDARD: PRT: 925 AA.  
 ID NRP2\_RAT  
 AC 035276;  
 DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE NEUROFILIN-2 PRECURSOR (VASCULAR ENDOTHELIAL CELL GROWTH FACTOR 165 RECEPTOR 2).  
 GN NRP2.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY;  
 RX MEDLINE=97433085; PubMed=9288754;  
 RA Kolodkin A.L., Levengood D.V., Rowe E.G., Tai Y.-T., Giger R.J., Ginty D.D.;  
 RA "Neurofilin is a semaphorin III receptor.";  
 RL Cell 90:753-762(1997).  
 CC -1- FUNCTION: HIGH AFFINITY RECEPTOR FOR SEMAPHORIN 3C, 3F, VEGF-165 AND VEGF-145 ISOFORMS OF VEGF, AND THE PLGF-2 ISOFORM OF PGF.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -1- TISSUE SPECIFICITY: FOUND IN CERTAIN NEURONAL POPULATIONS OF THE CNS AND IN OTHER NONNEURONAL TISSUES INCLUDING MESENCHYMAL TISSUE LINING IN THE RIBS.  
 CC -1- SIMILARITY: BELONGS TO THE NEUROFILIN FAMILY.  
 CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF016297; AAC53338.1; -;  
 DR InterPro: IPR00421; -;  
 DR InterPro: IPR00859; -;  
 DR InterPro: IPR00998; -;  
 DR Pfam: PF00431; CUB; 2.  
 DR Pfam: PF00629; MAM; 1.  
 DR Pfam: PF00754; F5\_F8\_type\_C; 2.  
 DR PROSITE: PS01180; CUB; 2.  
 DR PROSITE: PS01285; FA58C\_1; 2.  
 DR PROSITE: PS01286; FA58C\_2; 2.  
 DR PROSITE: PS50060; MAM\_2; 1.  
 KW Transmembrane; Glycoprotein; Neurone; Signal; Repeat; Receptor.  
 FT SIGNAL 1 22  
 FT CHAIN 23 925  
 FT DOMAIN 23 858  
 FT TRANSMEM 859 883  
 FT DOMAIN 884 925  
 FT DOMAIN 28 142  
 FT DOMAIN 149 267  
 FT DOMAIN 277 427  
 FT DOMAIN 434 592  
 FT DOMAIN 642 802  
 FT DISULFID 83 105  
 FT DISULFID 28 55  
 FT DISULFID 149 175  
 FT DISULFID 208 230  
 FT DISULFID 277 427  
 FT DISULFID 434 592  
 FT CARBOHYD 152 157  
 FT CARBOHYD 157 157  
 FT CARBOHYD 629 629  
 FT CARBOHYD 833 833

FT CARBOHYD 834 834 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 925 AA; 103896 MM; 3BF62903F644851C CRC64;  
 Query Match 5.9%; Score 464.5; DB 1; Length 925;  
 Best Local Similarity 28.6%; Pred. No. 1,2e-22;  
 Matches 167; Conservative 85; Mismatches 186; Indels 145; Gaps 33;  
 QY 967 DEFDCK-AMAYFSDVDLEK----DVHSGILGPLVCHTNTL-----NPAHGQVYQVEFA 1016  
 DB 79 EKHDCKYDFEIRHGDSESDLLGRKCGNAPRTIISGSLYIKFTSDYAKQA--GFS 136  
 QY 1017 LEFTIFDETKSWYFTENMERNCRAPCNIQMEDPTFEKYNRFPAINGYIMDTLPLGLVMAOD 1076  
 DB 137 LRYEIF-KTGS----EDCSKNFTSP--NGTIESGPFEKYP----- 170  
 QY 1077 QRIRWLLSGNSNENHSHIFSGHVTYVKKEEYKALNLYPGVEYEMLPKAG--- 1133  
 DB 171 -----HNLDCTETILAKPRMEIILQFL--TFD-LEHDPLOYEGED 207  
 QY 1134 -----IW-----RVECLTGEH-----LHAGMST-----LELY 1155  
 DB 208 CKYDMLDWDGIRHVGRLIKYGTGTPSKLNSGTILSTFHTDMAVAKDGFSAHYLY 267  
 QY 1156 YSN-----KQTPPLGMAAGHIRDQITASGY--GQMAPKLARLHYSGINAW-----ST 1203  
 DB 268 HQPPENFQCNAPLGMESGRIANEQISASTSFDGRWTFQGRHL--GDNGMTPFNVDN 325  
 QY 1204 KEFSPKTYDGLAPMIIHGIKTGA--RQKFSLEYISQFTIYSLDGKKWQYRGNSGT 1261  
 DB 326 KE--YLOVDLRLPLMTALATGALSTRETQGYVAKSTKLEVSTGEDMVMYRRGKNH- 381  
 QY 1262 LMYFGNVSSGIGKHNIFNPILARYIRLPHYISIRSTLRMLMGCDLNS--CSMPLGM 1319  
 DB 382 -KVFQANNDATELYLNKIKHTPLTRIRIRPQWHGIALRLIELGRCRTDAPCSNMLGM 440  
 QY 1320 ESKAISDAQITASSYFTNMFATWSPSKARLHLOGRSNAN---RPOVNNPKEMLYQVFOKT 1376  
 DB 441 LSGLIADTQISASS--TREY-LMSPSAARL-VSSRS-GWFPNPPQAQPEEWLYQVLDGTP 495  
 QY 1377 MKVTGYTIOGVK--SLLTSM----YVKEFLISSSDQGHQWTLFPQNGKV---KYFQGNOD 1427  
 DB 496 KTVKGVIIIGARGDSITLMEARAFYRKFRKVSISLGKQWE-YIDDPKIQOPKLFEGNMH 554  
 QY 1428 SFTPVVNSLDPPLRLTRYLRIRHPSQWVHQ--IALMEVYLGCEAD 1469  
 DB 555 YDTPDIRREE-PVPAQYVAVYPERMSPAIGIKRLVYLGCDMTD 596

Search completed: November 17, 2001, 13:24:36  
 Job time: 140 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2001, 13:22:11 ; Search time 26.31 Seconds  
(without alignments)  
4258.942 Million cell updates/sec

Title: US-09-689-430-2  
Perfect score: 7869  
Sequence: 1 MQLSTCFCLRLRFCSA.....WVHQLRMVELGCEAQLDY 1471

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | ID        | Description               |
|------------|--------|-------------|--------|-----------|---------------------------|
| 1          | 7412   | 94.2        | 2351   | 1 E2HU    | coagulation factor        |
| 2          | 6347   | 80.7        | 2133   | 2 T42763  | coagulation factor        |
| 3          | 6301   | 80.1        | 2319   | 2 A47004  | coagulation factor        |
| 4          | 2406.5 | 30.6        | 2224   | 1 KFH05   | coagulation factor        |
| 5          | 2394.5 | 30.4        | 2183   | 2 T42764  | coagulation factor        |
| 6          | 2379   | 30.2        | 2211   | 1 KFB05   | coagulation factor        |
| 7          | 1729   | 22.0        | 1069   | 1 KUH0    | coagulation factor        |
| 8          | 1681   | 21.4        | 1059   | 1 A35210  | ferroxidase (EC 1.11.1.1) |
| 9          | 1104   | 14.0        | 216    | 2 A44258  | ferroxidase (EC 1.11.1.1) |
| 10         | 663    | 8.3         | 427    | 2 JC4915  | factor VIII-associ        |
| 11         | 657    | 8.3         | 463    | 1 A36479  | agg protein precu         |
| 12         | 650    | 8.3         | 409    | 2 T11743  | milk fat globule m        |
| 13         | 635    | 8.1         | 401    | 2 S65138  | glycoprotein anti         |
| 14         | 635    | 8.1         | 427    | 2 S74211  | PAS-6/7 protein pr        |
| 15         | 543    | 6.9         | 869    | 2 A25945  | coagulation factor        |
| 16         | 443    | 5.6         | 927    | 1 J00948  | A5 antigen precurs        |
| 17         | 424.5  | 5.4         | 218    | 2 A47285  | milk fat globule p        |
| 18         | 306.5  | 3.9         | 313    | 2 S52093  | hemocytin - silkwo        |
| 19         | 262    | 3.3         | 845    | 2 JC5256  | antipocyte transcri       |
| 20         | 221    | 2.8         | 719    | 2 S11739  | transcription repr        |
| 21         | 211    | 2.7         | 1283   | 2 T113799 | neurexin IV - frui        |
| 22         | 207.5  | 2.6         | 737    | 2 T131349 | hypothetical prote        |
| 23         | 206.5  | 2.6         | 737    | 2 T15615  | hypothetical prote        |
| 24         | 193    | 2.5         | 913    | 2 A48280  | receptor tyrosine         |
| 25         | 192    | 2.4         | 876    | 2 A49508  | protein-tyrosine k        |
| 26         | 185.5  | 2.4         | 855    | 2 S42621  | protein-tyrosine k        |
| 27         | 185.5  | 2.4         | 910    | 2 A53137  | tyrosine kinase re        |
| 28         | 173.5  | 2.2         | 819    | 2 I48859  | tyro 10 receptor k        |
| 29         | 171    | 2.2         | 1381   | 2 T31083  | paranodin - rat           |

|    |       |     |      |          |                     |
|----|-------|-----|------|----------|---------------------|
| 30 | 171   | 2.2 | 1385 | 2 T14158 | neurexin IV - mous  |
| 31 | 150   | 1.9 | 1883 | 2 G82875 | hypothetical prote  |
| 32 | 149.5 | 1.9 | 578  | 2 S63533 | L-ascorbate oxidas  |
| 33 | 144.5 | 1.8 | 2925 | 2 T00133 | RNA-directed RNA p  |
| 34 | 142   | 1.8 | 622  | 2 S62580 | probable multiproc  |
| 35 | 141   | 1.8 | 1217 | 2 T18209 | mng protein - Bac   |
| 36 | 141   | 1.8 | 2401 | 2 T28676 | rhoptly protein -   |
| 37 | 138   | 1.8 | 1193 | 2 JC4901 | nontoxic-nonhemag   |
| 38 | 137.5 | 1.7 | 1178 | 2 S54073 | probable membrane   |
| 39 | 136.5 | 1.7 | 1196 | 2 S46430 | botulinum neurotox  |
| 40 | 135.5 | 1.7 | 1196 | 2 J01467 | toxlin, nontoxic co |
| 41 | 135   | 1.7 | 343  | 2 T35030 | probable copper ox  |
| 42 | 134.5 | 1.7 | 567  | 2 T44928 | L-ascorbate oxidas  |
| 43 | 134.5 | 1.7 | 791  | 2 T16031 | hypothetical prote  |
| 44 | 132.5 | 1.7 | 1132 | 2 H82887 | hypothetical prote  |
| 45 | 131.5 | 1.7 | 580  | 2 F84828 | probable laccase (  |

## ALIGNMENTS

RESULT 1  
E2HU  
coagulation factor VIII precursor [validated] - human  
N:Alternate names: antihemophilic factor A; coagulation factor VIIIc; prococagulant co  
C:Species: Homo sapiens (man)  
C>Date: 28-Aug-1985 #sequence\_revision 28-Aug-1985 #text\_change 08-Dec-2000  
C:Accession: I54318; A00525; I58059; A23584; A26174; A42348; A43986; S63527; S66445;  
R:Gitschler, J.; Wood, W.I.  
Hum. Mol. Genet. 1, 199-200, 1992  
A:Title: Sequence of the exon-containing regions of the human factor VIII gene.  
A:Reference number: I54318; MUID:93265012  
A:Accession: I54318  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1921, 'S', 1923-2351 <RES>  
A:Cross-references: GB:M8648; NID:g182381; PIDN:AAA52420.1; PID:g182383  
R:Wood, W.I.; Capon, D.J.; Simonsen, C.C.; Eaton, D.L.; Gitschler, J.; Keyt, B.; Seeb  
Nature 312, 330-337, 1984  
A:Title: Expression of active human factor VIII from recombinant DNA clones.  
A:Reference number: A00525; MUID:85061548  
A:Accession: A00525  
A:Molecule type: mRNA  
A:Residues: 1-2351 <MOO>  
A:Cross-references: EMBL:X01165; EMBL:X01166; EMBL:X01179  
R:Toole, J.C.; Knopf, J.L.; Wozney, J.M.; Sultzman, L.A.; Buecker, J.L.; Pittman, D.D  
s, D.N.; Hewick, R.M.  
Nature 312, 342-347, 1984  
A:Title: Molecular cloning of a cDNA encoding human antihemophilic factor.  
A:Reference number: I58059; MUID:85061550  
A:Accession: I58059  
A:Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-74, 'V', 76-1259, 'E', 1261-2351 <RE2>  
A:Cross-references: GB:K01740; NID:g182802; PIDN:AAA52484.1; PID:g182803  
R:Truett, M.A.; Blacher, R.; Burke, R.L.; Caput, D.; Chu, C.; Dhna, D.; Hattog, K.; K  
B.; Randojph, A.; Urdea, M.S.; Valenzuela, P.; Dahl, H.H.; Favalato, J.; Hansen, J.;  
DNA 4, 333-349, 1985  
A:Title: Characterization of the polypeptide composition of human factor VIII:C and t  
A:Reference number: A23584; MUID:86081164  
A:Accession: A23584  
A:Molecule type: mRNA  
A:Residues: 1-2351 <TRD>  
A:Cross-references: GB:M4113; NID:g182817; PIDN:AAA52485.1; PID:g182818  
R:Eaton, D.; Rodriguez, H.; Vohar, G.A.  
Biochemistry 25, 505-512, 1986  
A:Title: Proteolytic processing of human factor VIII. Correlation of specific cleavag  
ity.  
A:Reference number: A26174; MUID:86159740  
A:Accession: A26174  
A:Molecule type: protein  
A:Residues: 20-36;392-399, 'X', 401-402;1668-1678;1709-1722, 'D', 1723-1725;1741-1755 <EA  
R:Pittman, D.D.; Wang, J.H.; Kaufman, R.J.

Biochemistry 33:335-3375, 1992  
A:Title: Identification and functional importance of tyrosine sulfate residues within re  
A:Reference number: A42348; MUID:92207952  
A:Accession: A42348  
A:Molecule type: protein  
A:Residues: 20-336;356-371;392-408;582-594;1668-1669,'X',1671;1672-1692;1693-1709-17  
A:Experimental source: recombinant material from Chinese hamster ovary cells  
A>Note: sequence extracted from NCBI backbone and corrected to correspond with the publi  
R:Pay, P.J.; Smudzin, T.M.  
J. Biol. Chem. 264, 14005-14010, 1989  
A:Title: Intersubunit fluorescence energy transfer in human factor VIII.  
A:Reference number: A43986; MUID:89340500  
A:Accession: A43986  
A:Molecule type: protein  
A:Residues: 'X',517-523;1853-1860,'X',1862-1864,'X',1866 <RAY>  
R:LeYTE, A.; van Schlingel, H.B.; Niehrs, C.; Huttner, M.P.; Mertens, K. ;  
J. Biol. Chem. 266, 740-746, 1991  
A:Title: Sulfation of Tyr(1680) of human blood coagulation factor VIII is essential for  
A:Reference number: A56126; MUID:91093266  
A:Contents: annotation: sulfation  
R:Gitschler, J.; Wood, W.I.; Goralaka, T.M.; Wion, K.L.; Chen, E.Y.; Eaton, D.H.; Vehar,  
Nature 312, 326-330, 1984  
A:Title: Characterization of the human factor VIII gene.  
A:Reference number: A56196; MUID:85061547  
A:Contents: annotation: Introns  
R:McMullen, B.A.; Fujikawa, K.; Davie, E.W.; Hedner, U.; Ezban, M.  
Protein Sci. 4, 740-746, 1995  
A:Title: Locations of disulfide bonds and free cysteines in the heavy and light chains C  
A:Reference number: A56216; MUID:9538127  
A:Contents: annotation: disulfide bonds  
A>Note: 329-Cys, 711-Cys, and 2019-Cys were shown to have free sulfhydryls  
R:Kjalkhe, M.; Hedning, A.; Talbo, G.; Persson, E.; Thomsen, J.; Ezban, M.  
Eur. J. Biochem. 234, 773-779, 1995  
A:Title: Amino acid residues 721-729 are required for full factor VIII activity.  
A:Reference number: S63527; MUID:96163459  
A:Accession: S63527  
A:Molecule type: protein  
A:Residues: 733-752;753-759 <KJA>  
R:Lind, P.; Larsson, K.; Spita, J.; Sygöw-Baeckman, M.; Almstedt, A.; Gray, E.; Sandberg  
Eur. J. Biochem. 232, 19-27, 1995  
A:Title: Novel forms of B-domain-deleted recombinant factor VIII molecules. Construction  
A:Reference number: S66445; MUID:96048024  
A:Accession: S66445  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1668-1685 <LIN>  
C:Comment: Factor VIII is activated by factor Xa and thrombin, but prolonged exposure pr  
C:Genetics:  
A:Gene: GDB:F8C  
A:Cross-references: GDB:119124; OMIM:306700  
A:Map position: Xq28-Xq28  
A:Intons: 48/2; 89/1; 130/1; 201/1; 224/1; 263/1; 337/1; 424/2; 481/3; 513/1; 584/3; 63  
C:Function:  
A:Description: acts as a cofactor, with calcium and phospholipid, for the factor IXa pro  
A:Pathway: blood coagulation  
C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase  
C:Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A; pla  
F:1-19/Domin: signal sequence #status predicted <SIG>  
F:20-2351/Product: coagulation factor VIII #status experimental <MAT>  
F:20-740/Product: coagulation factor VIIIA heavy chain #status experimental <ACH>  
F:20-356/Domin: A1 <DA1>  
F:23-348/Domin: ferroxidase repeat homology <FO1>  
F:32-759/Domin: A2 <DA2>  
F:402-730/Domin: ferroxidase repeat homology <FO2>  
F:760-1667/Domin: B <DBO>  
F:1668-2351/Product: coagulation factor VIIIA light chain #status experimental <ACL>  
F:1709-2038/Domin: A3 <DA3>  
F:1716-2038/Domin: ferroxidase repeat homology <FO3>  
F:2039-2191/Domin: C1 <DC1>  
F:2039-2188/Domin: discoidin I amino-terminal homology <DN1>  
F:2192-2351/Domin: C2 <DC2>  
F:2192-2345/Domin: discoidin I amino-terminal homology <DN2>  
F:2192-2345/Domin: discoidin I amino-terminal homology <DN2>  
F:60;258,601,776,803,847,919,962,982,1020,1024,1074,1085,1204,1274,1278,1301,1319,1403,1

| Query Match           | 94.2%  | Score 7412   | DB 1       | Length 2351 |
|-----------------------|--|--------------|------------|-------------|
| Best Local Similarity | 62.5%  | Pred. No. 0  |            |             |
| Matches 1470          | Conservative 0   | Mismatches 1 | Indels 880 | Gaps 1      |
| Qy 1                  | MOEISTCEPFLCLRFCCSARTRYLAGVELSMWDMOSDLGELVDARFPRVKSPPEN        | 60           |            |             |
| Db 1                  | MOEISTCEPFLCLRFCCSARTRYLAGVELSMWDMOSDLGELVDARFPRVKSPPEN        | 60           |            |             |
| Qy 61                 | TSVYVKTLFEYFTVHLNLIAKPRPPWMLGPTIOAEVYDTVITLKNMASHPSLHAY        | 120          |            |             |
| Db 61                 | TSVYVKTLFEYFTVHLNLIAKPRPPWMLGPTIOAEVYDTVITLKNMASHPSLHAY        | 120          |            |             |
| Qy 121                | GVSYWKASGABAYDDQTSQREKEDDKVFPGGSHTYVMQYLKENGPMASDPLCLATYSLSH   | 180          |            |             |
| Db 121                | GVSYWKASGABAYDDQTSQREKEDDKVFPGGSHTYVMQYLKENGPMASDPLCLATYSLSH   | 180          |            |             |
| Qy 181                | VDLYVDLMSGTIGALLVCREGSLAKKETQTLHKFTLFAVDEKSMHSEFKNSLMDRD       | 240          |            |             |
| Db 181                | VDLYVDLMSGTIGALLVCREGSLAKKETQTLHKFTLFAVDEKSMHSEFKNSLMDRD       | 240          |            |             |
| Qy 241                | AASARAKPMKMTVNGVNRSLPLGLGCHARKSVYMHVITGMGTPPEVHSIFLEGHTFLVRNH  | 300          |            |             |
| Db 241                | AASARAKPMKMTVNGVNRSLPLGLGCHARKSVYMHVITGMGTPPEVHSIFLEGHTFLVRNH  | 300          |            |             |
| Qy 301                | ROASLEISPIITFLFQAOTLLMDLGOFLFCHISSHOHGMEAYVYVDSCEPEPQLRMKNNE   | 360          |            |             |
| Db 301                | ROASLEISPIITFLFQAOTLLMDLGOFLFCHISSHOHGMEAYVYVDSCEPEPQLRMKNNE   | 360          |            |             |
| Qy 361                | EAEDVDLDTSEMDVYRFPDDNSPFIQISVAKKHKKTWVHYIAAEEDMDYAPLVLA        | 420          |            |             |
| Db 361                | EAEDVDLDTSEMDVYRFPDDNSPFIQISVAKKHKKTWVHYIAAEEDMDYAPLVLA        | 420          |            |             |
| Qy 421                | PDDRSYKSOYLNNNGPQIRGKRYKKVRFMAYYDEFKTRREALQHESSGILGPLLYGEVDTL  | 480          |            |             |
| Db 421                | PDDRSYKSOYLNNNGPQIRGKRYKKVRFMAYYDEFKTRREALQHESSGILGPLLYGEVDTL  | 480          |            |             |
| Qy 481                | LIIFKQASRPYNYI PHGITVDVRLPYSLRLPEKGVKHLKDFPILDGELTFKMYKVTVEDGP | 540          |            |             |
| Db 481                | LIIFKQASRPYNYI PHGITVDVRLPYSLRLPEKGVKHLKDFPILDGELTFKMYKVTVEDGP | 540          |            |             |
| Qy 541                | TKSDPRCLITRYSSSVNNERDLASGLTGPLLICIKESYDQNGOIMSKRNVILFVSFDE     | 600          |            |             |
| Db 541                | TKSDPRCLITRYSSSVNNERDLASGLTGPLLICIKESYDQNGOIMSKRNVILFVSFDE     | 600          |            |             |
| Qy 601                | NRSWLTENIORFLCPNPAQVLEDEPEOASINIMHSINQYVDSIQLSCLHEAVAYWILS     | 660          |            |             |
| Db 601                | NRSWLTENIORFLCPNPAQVLEDEPEOASINIMHSINQYVDSIQLSCLHEAVAYWILS     | 660          |            |             |
| Qy 661                | IGAQDTFLVSFSGYTFKHKMYEDTTLTFPFGSETVFMSENPGMLTLCGNSDFNRNG       | 720          |            |             |
| Db 661                | IGAQDTFLVSFSGYTFKHKMYEDTTLTFPFGSETVFMSENPGMLTLCGNSDFNRNG       | 720          |            |             |
| Qy 721                | MTALLKSSCKKNQDGYIEDSYEDISATYLSLKNNAIEPRSSQNSRHPSTROKOFNAT      | 778          |            |             |
| Db 721                | MTALLKSSCKKNQDGYIEDSYEDISATYLSLKNNAIEPRSSQNSRHPSTROKOFNAT      | 778          |            |             |
| Qy 779                | -----  | 778          |            |             |
| Db 781                | PENDIEKTDPMFAHRTPMFKIQNVSSSDDLMLLRQSPTPHGLSLSDIQEAKYETFFSDPS   | 840          |            |             |
| Qy 779                | -----  | 778          |            |             |

```
Db 841 PGAIIDNNLSLSEMTHTFRPOLHHSGDWFTPEESGLQRLNEKLGTTAATFLKIDFKVSST 900
QY 779 ----- 778
Db 901 SNNLISTIPSDNLAAGTDNTSSIGPPSMFHVHDSQDITTLFGKSSPLTESGGLSLSEE 960
QY 779 ----- 778
Db 961 NNDKSLBESGLMNSQSSMGKVVSSTESGRLEKGRKRAHGPALLTDNALFKVISILKTN 1020
QY 779 ----- 778
Db 1021 KTSNNSATNRKTHIDPSLLIENSPVMQNTLESDETEKKYVPLIHDMRLMDKNATLRL 1080
QY 779 ----- 778
Db 1081 NMSNKTSSSKMEXVQCKEGRIPPDQNPDMSEFKMLFLPESARWIOFTHGKNSLNSG 1140
QY 779 ----- 778
Db 1141 QCPSPKOLVSLCPKESVSGQNFLEKKNKYVGKEFTKDVGLKEMVFPSSRMLFLTNDN 1200
QY 779 ----- 778
Db 1201 LHENNTHNOEKKIOBEIEKKETLIGENVYLPQIHTVTGFKNFMKMLFLSTRQNVESGYD 1260
QY 779 ----- 778
Db 1261 GAYAVLQDFRSLANDSTNRKTKHFAHFSKGEENLEGLNQTKOIVEKYACTRISBNT 1320
QY 779 ----- 778
Db 1321 SQQNVYTORSKRALKQFRLPLETELEKRIIYDITSTOMSKMKLTPSTLTQIDYNEKE 1380
QY 779 ----- 778
Db 1381 KGAITQPSLSDCLTRSHSIPQANRSPDLIAKVSFSPSIRPIYLTVLFDQNSSHLPAASY 1440
QY 779 ----- 778
Db 1441 RKKDSVOESSHFLOGAKKNLSLAILTLEMTGDQREVSLGTSATNSVTKVENTVLP 1500
QY 779 ----- 778
Db 1501 KPDLKTSKVELLPKVMHYQKDLFPPTETSNCSPGHLDLVBSGLQGTGGAIKMEANRP 1560
QY 779 ----- 778
Db 1561 GKVPFLVATESSAKTPSKLIDPLAMDNHYGTQIPKEEMKSGEKSPKTAFAKKDTIISL 1620
QY 779 ----- 778
Db 1621 NACESNHAIAINEGONRPEIEVTAQKQTERILCSQNPVYLKRHRQRETRTTLQSDQE 1680
QY 801 IDYDITIVEMKKEDFDIYDEDENSPRSFQKTRHYFLAVERLMDYGMSSSPVLRNR 860
Db 1661 IDYDITIVEMKKEDFDIYDEDENSPRSFQKTRHYFLAVERLMDYGMSSSPVLRNR 1740
QY 861 AQSQSVPOFKKVVFOETDGSFTQPLRYGELNEHLGLLPYTRAEVEDINIMVTFNQASR 920
Db 1741 AQSQSVPOFKKVVFOETDGSFTQPLRYGELNEHLGLLPYTRAEVEDINIMVTFNQASR 1800
QY 921 PYSFTSSLSISTEEDORQCAEPKKNVKNPNETKTYFWKYQHMAAPTKDEFDCKAMAYFSDV 980
Db 1801 PYSFTSSLSISTEEDORQCAEPKKNVKNPNETKTYFWKYQHMAAPTKDEFDCKAMAYFSDV 1860
QY 981 DLEKDVHSLGLGPLLVCHTNTLNPAGROVYQOEFALFETJEDETKSMVFTENMRNCRA 1040
Db 1861 DLEKDVHSLGLGPLLVCHTNTLNPAGROVYQOEFALFETJEDETKSMVFTENMRNCRA 1920
QY 1041 PCNIOMEDPTEKENYRFAHNGYIMDTLPLGLVMAODORIRWYLLSMGSNENIHSHFSGH 1100
```

```
Db 1921 PCNIOMEDPTEKENYRFAHNGYIMDTLPLGLVMAODORIRWYLLSMGSNENIHSHFSGH 1980
QY 1101 VFTVARKKEEYKMAALNLYGVFEYTEMELPSKAGIMRVBELIEHLAGSTLFLYYSNKC 1160
Db 1981 VFTVARKKEEYKMAALNLYGVFEYTEMELPSKAGIMRVBELIEHLAGSTLFLYYSNKC 2040
QY 1161 QPPLGMASGHINDFOITASGOYGQWAPKLARLHYSGSINAMSTKEPFSWIKYDLAPMTI 1220
Db 2041 QPPLGMASGHINDFOITASGOYGQWAPKLARLHYSGSINAMSTKEPFSWIKYDLAPMTI 2100
QY 1221 HGIKTQGARQKPESSLYISQFIIMYSLDGKKWQYRKNSTGLTMVEFGNVSSGIRKHNFN 1280
Db 2101 HGIKTQGARQKPESSLYISQFIIMYSLDGKKWQYRKNSTGLTMVEFGNVSSGIRKHNFN 2160
QY 1281 PPIIARYIRLHPTHTYSINSTLMELMCDLNSCSPMLGESAISDAQITASSYFTNMFA 1340
Db 2161 PPIIARYIRLHPTHTYSINSTLMELMCDLNSCSPMLGESAISDAQITASSYFTNMFA 2220
QY 1341 TWPSPKARLHLOGRSNAMPQVNNPKEMLOVDFQOTMKVTGVTGQVSKLTSMTVKEFL 1400
Db 2221 TWPSPKARLHLOGRSNAMPQVNNPKEMLOVDFQOTMKVTGVTGQVSKLTSMTVKEFL 2280
QY 1401 ISSSODGHQWTLFPQNGKVKVEQGNDSFTPVVNSLDPELTLRYLRIRHPQSWHOIALRM 1460
Db 2281 ISSSODGHQWTLFPQNGKVKVEQGNDSFTPVVNSLDPELTLRYLRIRHPQSWHOIALRM 2340
QY 1461 EYLGCEADPLY 1471
Db 2341 EYLGCEADPLY 2351
```

## RESULT 2

```
142763
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jun-2000
C:Accession: T42763
R:collar: P.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z2269
A:Accession: T42763
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2133 <IOL>
A:Cross-references: EMBL:U49517; NID:q1511633; PID:q1511634; PIDN:AB06705.1
C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxid
C:Keywords: acute phase; blood coagulation; duplication; 9lycoprotein; hemophilia A;
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-2133/Product: coagulation factor VIII #status predicted <MAT>
F:23-349/Domain: ferroxidase repeat homology <FOX1>
F:402-730/Domain: ferroxidase repeat homology <FOX2>
F:1498-1820/Domain: ferroxidase repeat homology <FOX3>
```

Query Match 80.7%; Score 6347; DB 2; Length 2133;

Best Local Similarity 58.0%; Pred. No. 0; Mismatches 126; Indels 664; Gaps 4;

Matches 1237; Conservative 107; Mismatches 126; Indels 664; Gaps 4;

```
QY 1 MQELSTCFELCLLRFCSATRRRYLAGAVELSDWYQSD-IGELPVDARFPFRPKSPPE 59
Db 1 MQELSTCFELCLLRFCSATRRRYLAGAVELSDWYQSD-IGELPVDARFPFRPKSPPE 60
QY 60 NTSVYKKTTLVEFTVHLFNIAKPRPPMGLIGPTIQAEVYDVTVYITLKNMASHVSLHA 119
Db 61 GPSVLYKKTVEFEFTDOLFVSARPRPMMGLIGPTIQAEVYDVTVYITLKNMASHVSLHA 120
QY 120 VGVSYKASSEGAEYDDQSORKEEDKVEFGGSHYVQVYKENGPMASDPLCLTYSLYS 179
Db 121 VGVSYKASSEGAEYDDQSORKEEDKVEFGGSHYVQVYKENGPMASDPLCLTYSLYS 180
QY 180 HVDLVKDNSGLIGALLVCREGLSLAKERTQTLHKFILLFAVFEDECKSWHSETKNSIAMDR 239
Db 181 HVDLVKDNSGLIGALLVCREGLSLAKERTQTLHKFILLFAVFEDECKSWHSETKNSIAMDR 240
```

|    |      |   |      |
|----|------|---|------|
| QY | 240  | DASASAMPMKHTVNGVNSLPLGLICGCHKSVWHVIGMGTPEVHSIFLEGHFTLVRN      | 299  |
| Db | 241  | DPAAPRQPMHTVNGVNSLSPGLICGCHKSVWHVIGMGTPEVHSIFLEGHFTLVRRH      | 300  |
| QY | 300  | HROASLEISPTFLTAQTLMLDGLQFLFLFCHIISSHQDGMCAVYKVSCEPDELRMKN     | 359  |
| Db | 301  | HROASLEISPTFLTAQTLMLDGLQFLFLFCHIISSHGHGMEAHVRVSCAEERQLRRKAD   | 360  |
| QY | 360  | EEAEYDDDLTDEBMDVYRFDDDNSSPFIQIRSAKKHPRKTWHYIAAEEEDMDYAPLV     | 419  |
| Db | 361  | EE-EDYDDNLTDDMDVRLDGDVRFQIRSAKKHPRKTWHYIAAEEEDMDYAPAVP        | 419  |
| QY | 420  | APDRASYKSOYLNNGPORIGRKYKKVRFMAVTDFTFKTREAIQHESSILGLLGEVDT     | 479  |
| Db | 420  | SPEDRKYKSLYLNSGPORTIRKRYKKARFVATVDTFTKRAIPRESGILGLLGEVGD      | 479  |
| QY | 480  | LIIIFKNQASRPYNIYPHGITDVRLPYSRRLPKGVKHLKOPRLIPGETFKYKMTVYEDG   | 539  |
| Db | 480  | LIIIFKNKASRPNIYPHGITDVRSALHPGRLLKGMKHLKMPRLIPGETFKYKMTVYEDG   | 539  |
| QY | 540  | PTKSDRCLTRYYSFVNMRDLASGLIGPLLCYKESVORGNQIMDKRNVLLFVFP         | 599  |
| Db | 540  | PTKSDRCLTRYTSSTINLEKDLASGLIGPLLCYKESVDQRGNQIMDKRNVLLFVFP      | 599  |
| QY | 600  | ENRSWYLTENIQFLPNPAGVQLEDPFOASNIMHSINGVYFDSLQSVCLHEVAYVYL      | 659  |
| Db | 600  | ENQSWYLAENIQFLPNPDGLQDPDPFOASNIMHSINGVYFDSLQSVCLHEVAYVYL      | 659  |
| QY | 660  | STDAQDFLSVFFSGTTFKKRKNYEDTLTFPFSGETVFNMSMKNPGMILGCHNSDRNR     | 719  |
| Db | 660  | SVQAQDFLSVFFSGTTFKKRKNYEDTLTFPFSGETVFNMSMKNPGMVLGCHNSDRNR     | 719  |
| QY | 720  | GMMALLKVSQCDNKNMGYEDSYEDISAYLLSKNNALIEPFSQNSNHPSTRKOF----     | 775  |
| Db | 720  | GMMALLKVSQCDNDIGDYDNTYEDIPGFLLSGKNVIEPFSFAQNSRPPASQKOFOTIT    | 779  |
| QY | 776  | -----   | 775  |
| Db | 780  | SPEDVDLPQSGERTQALEELSVPSGSGMLLGQNAPRHGSSSSDLQEARNEADYLPG      | 839  |
| QY | 776  | -----   | 775  |
| Db | 840  | AREBMTAPSAARLRPELHHSABERYLTPEPEKELKLDKSMSSSSDLKTSPITSPDLS     | 899  |
| QY | 776  | -----   | 775  |
| Db | 900  | AETERTHSLGPRPHQVNFRSQLAIVLGKNSHPIGAGVPLGSTEEDEHSSLGENVSPVE    | 959  |
| QY | 776  | -----   | 775  |
| Db | 960  | SDGIFEKERAHGASLT KDVLEKVNISLVKTKARVYLTKNRKIHIDDAALLTENRASA    | 1019 |
| QY | 776  | -----   | 775  |
| Db | 1080 | AVEEBEELSKGKEMMLPNSELFTLNSADVOGNDTHSQGSKSREEMERREKLVOEKVDLPQ  | 1139 |
| QY | 776  | -----   | 775  |
| Db | 1140 | VYTAGTKNLFNIHFHOSTEPSEVEGEDGSHAPVPODSSRLNDSABEAETHIAHFAISIRE  | 1199 |
| QY | 776  | -----   | 778  |
| Db | 1200 | EAPLEAPGNKTGSPRSAPVRVKQSLKQIRLPLEEIKPERGVYLANFTSTRMSSSPILQ    | 1259 |
| QY | 779  | -----   | 778  |
| Db | 1260 | GAKRNMLSPFLTLEMAAGGOKISALGKSAAGPILASGLKEKAVLSSAGLSEASGAKAEFLP | 1319 |

|    |      |   |      |
|----|------|---|------|
| QY | 779  | -----   | 778  |
| Db | 1320 | KVRVHREDLLPQKTSNVSCHAGLGEIEFLQKTRGVNLTKNVNRPERTSKLLGPRMKE       | 1379 |
| QY | 779  | -----   | 778  |
| Db | 1380 | WESLEKSPKSTALTRKDIISLPLDRHESNHSIAKNDCQAEQTOREAAMTKOGPGLCAP      | 1439 |
| QY | 779  | -PPLKAKHOREIRTTYLQSQOEELDYDDTISVMKKEEDDIYDEDEDNOSPFSOKKTRHY     | 837  |
| Db | 1440 | KPYLRKHQRDILSPTFQPEBEMDDDDIFSTFKEDIEDIDIGEDENDPSPFKKTRHY        | 1499 |
| QY | 838  | FIAVERLMDYGMSSSPHVLNRNAGSGSVPOFKVVFQOEFTGSGFPQYRRELENIHLG       | 897  |
| Db | 1500 | FIAVAEDLMDYGMSESPRALRNRAONGEVRFKVVFRFPAQSGSPQSPYRBEUNKHIGL      | 1559 |
| QY | 898  | LGPYIRAEVDNIMVTFPRNAPSRYFSLSJYEEEDQOGAPPRKAPVETKTYRKK           | 957  |
| Db | 1560 | LGPYIRAEVDNIMVTFKKNASPRYSLSJYPPDOOGAEPRIINFQAPRETRYEAK          | 1619 |
| QY | 958  | VOHMAPTKDEPQCKAMAYSDVDLEKDVHSGILGPLVCHNTPLNPAHGRQVTVQEFAL       | 1017 |
| Db | 1620 | VOHMAPTKDEPQCKAMAYSDVDLEKDVHSGILGPLICRANTPLNAAHGRQVTVQEFAL      | 1679 |
| QY | 1018 | FTTIFDEKTSWYFTENMERNCRAPCNLOMEDPTFKENYRPHALINGYIMDTLGLVMAQOQ    | 1077 |
| Db | 1680 | FTTIFDEKTSWYFTENYERNCRAPCHLOMEDPPLKENYRPHALINGYIMDTLGLVMAQNO    | 1739 |
| QY | 1078 | RIRWYLLSMGSNNHNSIHRSGHVFTYRKKKEEKMAVLYLYGVPEYEMLSKAGIMRY        | 1137 |
| Db | 1740 | RIRWYLLSMGSNNHNSIHRSGHVFSVRKKKEEKMAVLYLYGVPEYEMLSKAGIMRI        | 1799 |
| QY | 1138 | ECLIGELHLAGMSTFLVYSNKCOTPLGMAHGHIRPQITAGSGQVGOMAPKLARLHSGS      | 1197 |
| Db | 1800 | ECLIGELHLAGMSTFTLYVSKECQAPLGMASGHIRPQITAGSGQVGOMAPKLARLHSGS     | 1859 |
| QY | 1198 | INAMSTKEPSPWIKVDLLAPMITHGIRKTOGAROKESSLYISQFTIMYSLDGKKMOTYKCN   | 1257 |
| Db | 1860 | INAMSTKDPHSMWIKVDLLAPMILHIGIMTOGAROKESLSLYISOFIIMYSLDGKRMOSYKCN | 1919 |
| QY | 1258 | STGTLWVFGNVDSGKHNIFNPPIIARYITRLHPHYISIRSLRMELGCOLNCSMPL         | 1317 |
| Db | 1920 | STGTLWVFGNVDSGKHNIFNPPIVARYIRLHPHYISIRSLRMELGCOLNCSMPL          | 1979 |
| QY | 1318 | GMEKASIAQOITASSYFTNMFATWSPSKARLHLQGRSNAMRPQVNNPKEMLOVDFOKTM     | 1377 |
| Db | 1980 | GMOKKAISQOITASSHLSISIFATWSPSOARLHLQGRTNAMRPVSSAEBEMLOVDFOKTV    | 2039 |
| QY | 1378 | KYNGVTFQGVKSLTLTSMYKVEPLISSQDCHQWTLFPOKGVKYNVQGNODSTFPVNSLD     | 1437 |
| Db | 2040 | KYTGITTOGVKSLTSSYKVEPLVSSQDERRKWTLLQGHKTRKVFQGNODSTFPVNAJD      | 2099 |
| QY | 1438 | PPLTRYLRITHPOSWHQIALRMEVLGCENADOLY                              | 1471 |
| Db | 2100 | PPLTRYLRITHPTSWAQHIALRLEVLCGERADOLY                             | 2133 |

|                     |  |               |
|---------------------|--|---------------|
| RESULT              | 3  |               |
| A47004              | coagulation factor VIII precursor - mouse                                |               |
| C:Species:          | Mus musculus (house mouse)   |               |
| C:Date:             | 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 18-Jun-1999      |               |
| C:Accession:        | A47004   |               |
| R:Elider, B.;       | Lakich, D.;  | Gitschier, J. |
| Genomics            | 16, 374-379, 1993  |               |
| A:Title:            | Sequence of the murine factor VIII cDNA.                                 |               |
| A:Reference number: | A47004; MUID:93300511  |               |
| A:Accession:        | A47004   |               |
| A>Status:           | preliminary  |               |
| A:Molecule type:    | mRNA   |               |
| A:Residues:         | 1-2319 <EID>   |               |
| C:Superfamily:      | coagulation factor VIII; discolidin I amino-terminal homology; ferrooxid |               |

RESULT 3  
A47004  
coagulation factor VIII precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 18-Jun-1999  
C:Accession: A47004  
R:Elder, B.; Lalkich, D.; Gitschier, J.  
Genomics 16, 374-379, 1993  
A:Title: Sequence of the murine factor VIII cDNA.  
A:Reference number: A47004; MUID:93300511  
A:Accession: A47004  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-2319 <ELD>  
A:Cross-references: GB:LO5573; NID:g192456; PIND:AAA37385.1; PID:g192457  
A:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxiid



F:1-19/Domain: signal sequence #status predicted <SIG>  
F:23-349/Domain: ferroxidase repeat homology <FO1>  
F:402-730/Domain: ferroxidase repeat homology <FO2>  
F:1686-2006/Domain: ferroxidase repeat homology <FO3>  
F:2007-2156/Domain: discoidin I amino-terminal homology <DN1>  
F:2160-2313/Domain: discoidin I amino-terminal homology <DN2>

Query Match 80.1%; Score 6301; DB 2; Length 2319;  
Best local similarity 53.8%; Pred. No. 0;  
Matches 1251; Conservative 102; Mismatches 111; Indels 862; Gaps 10;

```

QY 1 MOEISTCFECLLRFCSCATRRYYLGAVALSMDWQSD-IGELVDAFRPPRVKSEPF 59
DB 1 MOIALFAPEFLSLFNFCSAIRYYLGAVALSMDWQSD-IGELVDAFRPPRVKSEPF 60
QY 60 NTSVYKKTLEVEVTHLFINAKPRPPMGLGPTIQAEVYDTVVITLKNASHPVSLHA 119
DB 61 NTSIYKKTVEVEYKDFLNIAKPRPPMGLGPTIMTEVHDVIVITLKNASHPVSLHA 120
QY 120 VGVSYWKASGAEVDDQTSQREKEDKVPFGSHYVWQVLKENGPMASDPLCLTYSL 179
DB 121 VGVSYWKASGAEVDDQTSQREKEDKVPFGSHYVWQVLKENGPMASDPLCLTYSL 180
QY 180 HVDLYKDLNSGLIGALLVCRGSLAKEKQTLLHKLFFLFAVFEDEKSNHSETKNSLMQDR 239
DB 181 HVDLYKDLNSGLIGALLVCRGSLAKEKQTLLHKLFFLFAVFEDEKSNHSETKNSLMQDR 240
QY 240 DAASARAPKMHVNGVYVNRSLPGLIGCHRSVYWHVIGMGTTPVHSLFLEGHTFLVRN 299
DB 241 DSASARAPKMHVNGVYVNRSLPGLIGCHRSVYWHVIGMGTTPHSLFLEGHTFLVRN 300
QY 300 HRQASLESPITFTLAOTLMDLGOFLFCHTSSHOHDMEAHYVAVDSCPPEPOLRMK-N 358
DB 301 HRQASLESPITFTLAOTLMDLGOFLFCHTSSHOHDMEAHYVAVDSCPPEPOLRMK-N 360
QY 359 NEEADYDDDDLTDEMDVVRFDSDNSPSFIOIRSAVKKHKPTWVYIAAEEEDMDYAPLV 418
DB 361 NEEADYDDDDLTDEMDVVRFDSDNSPSFIOIRSAVKKHKPTWVYIAAEEEDMDYAPLV 418
QY 419 LAPDRRSKSYQLNNGPQIRGIRKKYKRVFMAVYDTEFTREAIQIESGILGPLYGEVD 478
DB 419 PTDSDGSKSYQLNNGPQIRGIRKKYKRVFMAVYDTEFTREAIQIESGILGPLYGEVD 478
QY 479 TLLIIFKNOARPNYVYHGTIDVRLYSRLPKVKHKLKOPFLPGELFKYKWTYVED 538
DB 479 TLLIIFKNOARPNYVYHGTIDVRLYSRLPKVKHKLKOPFLPGELFKYKWTYVED 538
QY 539 GPTKSDPCLTRYSSSFVNMERDLASGLIGPLLCYKESVDQRGNOIMSDKRNVLFSVF 598
DB 539 GPTKSDPCLTRYSSSFVNMERDLASGLIGPLLCYKESVDQRGNOIMSDKRNVLFSVF 598
QY 599 DENRSMVTENIQRLPNPAGVQLEDPFOASNMHSINGVYVDSLOLSVCLHEVAYWI 658
DB 599 DENRSMVTENIQRLPNPAGVQLEDPFOASNMHSINGVYVDSLOLSVCLHEVAYWI 658
QY 659 LSIQAQDFELSVFSGYFFKHKMYEDTLTPPSGGEVFMSEMPGLMILGCHNSDRN 718
DB 659 LSVGAQDFELSVFSGYFFKHKMYEDTLTPPSGGEVFMSEMPGLMILGCHNSDRN 718
QY 719 RGMATLLKVVSSGDKNTGYDYDSYEDISAVLLSKNNAIEPFSFSONSHPTROKOF--- 775
DB 719 RGMATLLKVVSSGDKNTGYDYDSYEDISAVLLSKNNAIEPFSFSONSHPTROKOF--- 775
QY 776 --- 775
DB 776 --- 775
QY 779 TIRPKNMEKIEPOEEELAEMLKVQSVSDMLLGGSHPRHGLFLSDGOEAIYEALHD 838
DB 779 TIRPKNMEKIEPOEEELAEMLKVQSVSDMLLGGSHPRHGLFLSDGOEAIYEALHD 838
QY 839 DHSFPAIDNSGSPSKVYQOLRPESHSEKIVTFPOGLOLRSNKSLLETTIEVKKMKLGLOV 898
DB 839 DHSFPAIDNSGSPSKVYQOLRPESHSEKIVTFPOGLOLRSNKSLLETTIEVKKMKLGLOV 898
QY 776 --- 775
```

```

DB 899 SSLPSNLMTTITLSDNLKATFEKTTDSGPPDMPVHSSSKLSTTARCAKAYSLVGSHVPLN 958
QY 776 --- 775
DB 959 AASENSDNLSDSTLMYSQESLPRDNLISLENDRLREKRFHIALTKDNTLFRKDNVSL 1018
QY 776 --- 775
DB 1019 MKNKTYNHSTTNEKLHESPTSIENSTTDDADAILKYNSEIQEVYALIHDTLLGKNST 1078
QY 776 --- 775
DB 1079 YLRLLHMLNRTTSTKNKIDIFHRKDEDPIDODENTIMPSKMLFLSESSNMFKKTGNNS 1138
QY 776 --- 775
DB 1139 LNSEOHSPPKQOLVIMFKRYVKNQSFLEKNKVTVEODGFTKNIGLKDMAFPHNSIFLT 1198
QY 776 --- 775
DB 1199 TLSNVHENGHRNOEKNIOEELKEALIEKVVLPVYHEATGSKNFLKDLILGTRONISL 1258
QY 776 --- 778
DB 1259 YEAVHPVLQNTISINNSTNTVQIHMEHFFKRRKKDETNSEGLVKNKTRKMYKVPSSQKNT 1318
QY 779 --- 778
DB 1319 TQSRKALGQFLSTQWLKTINCSTQCIHQIDHSKEMKFTYKSLSDSSVYKSTTQTN 1378
QY 779 --- 778
DB 1379 SSDSHIVTSAPPIDLRSPQNKFSHYQASSYIYDEKTKSRIOESNLFKETKINNP 1438
QY 779 --- 778
DB 1439 SLAILPMMKFIIDQGFSTSPGKSNSTSVYKRENIIFLKPITPESGKIELLPQYSIOQE 1498
QY 779 --- 778
DB 1499 EILPETSHSGPGLNLKKEVLAQIQGPTKWKAKRHGESIKGTESKNTRSKXLNH 1558
QY 779 --- 778
DB 1559 AMDHYAAOIPKDMKSKSEKSPETIISIKOEDTILSLRPHGNSHIGANEKQMPQRETTW 1618
QY 779 --- 825
DB 1619 VKQGTQRTCSQIRPVLRKHQREL--SAFQSQEATVDQDATIE-TIEDPDYSEDIKQ 1675
QY 826 SPRSFQKTRHYFLAVERLMDYGNSSPHVLRNAGSGSVPOFKVVFQEFDDSGFTOP 885
DB 1676 GPRSFQKTRHYFLAVERLMDYGNSSPHVLRNAGSGSVPOFKVVFQEFDDSGFTOP 1734
QY 886 LYRGLNLHGLIGPYIAAEVEDNIMTFERNQASRPYSFYSLLSYEEDQAGAPRRNF 945
DB 1735 LYRGLNLHGLIGPYIAAEVEDNIMTFERNQASRPYSFYSLLSYEEDQAGAPRRNF 1793
QY 946 VKPNETKTYFMKVQHMAAPTDEDFCKAMAYFSVDLKKDVHSGILGLVLCHTNTLPA 1005
DB 1794 VKPNETKTYFMKVQHMAAPTDEDFCKAMAYFSVDLKKDVHSGILGLVLCHTNTLPA 1853
QY 1006 HGRQVTVQEFALFTIFDETKSWYFTENMERNCRAPCNIQMEDPTKENYRPHALINGYIM 1065
DB 1854 HGRQVTVQEFALFTIFDETKSWYFTENMERNCRAPCNIQMEDPTKENYRPHALINGYIM 1913
QY 1066 DTLGGLVMAODORTIRWYLLSGNSNENIHSIHSGHVFYVRKKEEYKMLVNLXPVFEFTY 1125
DB 1914 DTLGGLVMAODORTIRWYLLSGNSNENIHSIHSGHVFYVRKKEEYKMLVNLXPVFEFTY 1973
QY 1126 EMLPSKAGIWRVECLIGELHLAGMSTLFLVSNKQCPFLGNASGHIRDFQTLASGOYQOW 1185
DB 1126 EMLPSKAGIWRVECLIGELHLAGMSTLFLVSNKQCPFLGNASGHIRDFQTLASGOYQOW 1185
```

Db 1974 EMIPBAGIWEVCELIQHEHQAQGSTFLVYSQCOIPLGMASTGTRDQITASGHYQW 2033  
Qy 1186 APKLARLHSGSINAMSTKEPFSWIKVDLAPMIHQIKTOGAROKFSSLYISOFITMYS 1245  
Db 2034 APNLRALHSGSINAMSTKEPFSWIKVDLAPMIHQIKTOGAROKFSSLYISOFITMYS 2093  
Qy 1246 LDGKRWQYIRGNSGTGLMWVFGNVDSSGKHNIFNPPIIARYIRLPHTHYSIRSTLMEL 1305  
Db 2094 LDGKRWQYIRGNSGTGLMWVFGNVDSSGKHNIFNPPIIARYIRLPHTHYSIRSTLMEL 2153  
Qy 1306 MCCDINSCGPMGMEKATSDAQITASSTFTNMFATWSPKARLHOGSRNARPOVNNP 1365  
Db 2154 MCCDINSCGPMGMEKATSDAQITASSTFTNMFATWSPKARLHOGSRNARPOVNNP 2213  
Qy 1366 KEMLOVDFOKTKMYGVTTQGVKSLTSMYKVEFLLISSQDGHQWTLFQNGKVKVFOGN 1425  
Db 2214 KEMLOVDFOKTKMYGVTTQGVKSLTSMYKVEFLLISSQDGHQWTLFQNGKVKVFOGN 2273  
Qy 1426 QDSFTPVVNSLDPRPLTRYLRIHPQSWHQAIALRMEVLGCEADOLY 1471  
Db 2274 QDSFTPVVNSLDPRPLTRYLRIHPQSWHQAIALRMEVLGCEADOLY 2319

RESULT 4

KRF05  
N:Alternate names: coagulation labile factor; proaccelerin  
C:Species: Homo sapiens (man)  
C:Date: 19-May-1989 #sequence revision 02-Jun-1995 #text\_change 08-Dec-2000  
C:Accession: A561172; A42344; A28028; A27498; A25897  
R:Clupe, L.D.; Moore, K.D.; Kane, W.H.  
Biochemistry 31, 3777-3785, 1992  
A:Title: Structure of the gene for human coagulation factor V.  
A:Reference number: A42344; MUID:92232668  
A:Accession: A561172  
A:Molecule type: DNA  
A:Residues: 1-2224 <CRI>  
A:Cross-references: GB:J05368  
A:Accession: A42344  
A:Molecule type: DNA  
A:Residues: 48-58;79-89;120-130;191-201;239-249;313-323;368-378;428-437;461-471;533-542;  
2070;2111-2120;2172-2181 <CR>  
R:Jenny, R.J.; Pltman, D.D.; Toole, J.J.; Kriz, R.W.; Aldape, R.A.; Hewick, R.M.; Kaufm  
Proc. Natl. Acad. Sci. U.S.A. 84, 4846-4850, 1987  
A:Title: Complete cDNA and derived amino acid sequence of human factor V.  
A:Reference number: A28028; MUID:87260886  
A:Accession: A28028  
A:Molecule type: mRNA  
A:Residues: 1-857, 'R', 859-864, 'R', 866-924, 'E', 926-1763, 'T', 1765-2212, 'T', 2214-2224 <JEN>  
A:Cross-references: GB:M16967  
A:Note: parts of this sequence, including the amino end of the mature protein, were deter  
R:kane, W.H.; Ichinose, A.; Hagen, F.S.; Davie, E.W.  
Biochemistry 26, 6508-6514, 1987  
A:Title: Cloning of cDNAs coding for the heavy chain region and connecting region of hum  
A:Reference number: A27498; MUID:88107560  
A:Accession: A27498  
A:Molecule type: mRNA  
A:Residues: 1-1284, 'I', 1286-1600 <KAN>  
A:Cross-references: GB:M1785  
A:Note: parts of this sequence were determined by protein sequencing  
R:kane, W.H.; Davie, E.W.  
Proc. Natl. Acad. Sci. U.S.A. 83, 6800-6804, 1986  
A:Title: Cloning of a cDNA coding for human factor V, a blood coagulation factor homolog  
A:Reference number: A25897; MUID:86313665  
A:Accession: A25897  
A:Molecule type: mRNA  
A:Residues: 1188-1215, 1315-2224 <KA2>  
A:Cross-references: GB:M14335  
A:Note: parts of this sequence were determined by protein sequencing  
R:keller, F.G.; Otefl, T.L.; Quinn-Allen, M.A.; Kane, W.H.  
Biochemistry 34, 4118-4124, 1995  
A:Title: Thrombin-catalyzed activation of recombinant human factor V.  
A:Reference number: A56139; MUID:95210278

A:Contents: annotation; thrombin cleavage sites  
C:Comment: Factor V is activated by thrombin and partially by coagulation factor Xa.  
C:Genetics:  
A:Gene: GDB:F5  
A:Cross-references: GDB:119896; OMIM:227400  
A:Map position: 1q23-1q23  
A:Intons: 53/2; 84/1; 125/1; 196/1; 244/1; 318/1; 373/2; 432/3; 466/1; 537/3; 588/1;  
C:Function:  
A:Description: acts as a cofactor, with calcium and phospholipid, for the factor Xa p  
A:Pathway: blood coagulation  
C:Superfamily: coagulation factor V; discoidin I amino-terminal homology; ferroxidase  
C:Keywords: blood coagulation; duplication; glycoprotein; phospholipid binding; plasm  
F:1-28/Domain: signal sequence #status predicted <Sig>  
F:29-2224/Product: coagulation factor V #status predicted <Mat>  
F:29-737/Product: coagulation factor Va heavy chain #status experimental <VAH>  
F:29-345/Domain: A1 <DA1>  
F:33-329/Domain: ferroxidase repeat homology <FO1>  
F:346-691/Domain: A2 <DA2>  
F:351-684/Domain: ferroxidase repeat homology <FO2>  
F:592-1573/Domain: B <DOB>  
F:1183-1461/Region: 9-residue repeats (Q-X-T/N-L-S-P-D-L-S)  
F:1574-2224/Product: coagulation factor Va light chain #status experimental <VAL>  
F:1574-1905/Domain: A3 <DA3>  
F:1581-1905/Domain: ferroxidase repeat homology <FO3>  
F:1667-1765/Region: phospholipid binding #status predicted  
F:1906-2064/Domain: C1 <DC1>  
F:1906-2064/Domain: discoidin I amino-terminal homology <DNI>  
F:2065-2224/Domain: C2 <DC2>  
F:2065-2221/Domain: discoidin I amino-terminal homology <DN2>  
F:51,55,229,297,460,468,554,741,752,760,776,782,821,938,977,1074,1083,1103,1106,1479,  
F:167-193,248-329,500-526,603-684,1125-1751,1907-2061,2066-2221/Disulfide bonds: #sta  
F:334-335/Cleavage site: Arg-Asn (protein A) #status predicted  
F:363,693,1546/Binding site: sulfate (Tyr) (covalent) #status predicted  
F:376-377/Cleavage site: Arg-Ser (coagulation factor Xa) #status predicted  
F:382,1338/Binding site: carboxylate (Asn) (covalent) #status absent  
F:534-535/Cleavage site: Arg-Gly (protein C) #status predicted  
F:737-738/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experiment  
F:1046-1047/Cleavage site: Arg-Thr (coagulation factor Xa, thrombin) #status experime  
F:1573-1574/Cleavage site: Arg-Ser (thrombin) #status experimental

Query Match 30.6%; Score 2406.5; DB 1; Length 2224;  
Best Local Similarity 26.3%; Pred. No. 5,4e-155;  
Matches 599; Conservative 279; Mismatches 489; Indels 909; Gaps 37;  
Qy 22 RRYVLGAVELSDMPQSDGELPVDARPPRVKSPFNTSVV-YKKTLEFVYHLFNI 80  
Db 32 KQYVAAQGISWYRE-----PTNSLNLSTSRKTIYRREYR-FKK 75  
Qy 81 AKRPRPMGLLPTIQAEYDVTIVTLKMAASHVSLAVGVSYKASGAEYDQTSOR 140  
Db 76 EKQSTISGLLPTQLAEVGDILKVFHKKADKPLSIHQGIRYSKLSGASYLDPFPA 135  
Qy 141 EKEDKVPFGSGSTTYWQYLKENGPMASDPLCLTYLSEHVDLVDNLGGLGALLVRE 200  
Db 136 EKKMDAVAPRGRETYEWSISEDSPTPHDRPCLTHLYYHNMLEDPSNLGLPLDICK 195  
Qy 201 GSLAKKETO-TLHK-FILFAVPEDEKSWHSETKNSLMODRDAASARAPKMHNTGVYN 258  
Db 196 GTLTGCTQKTPKQYLVLFVAFDESKWSQSS-----LMTYNGVYN 239  
Qy 259 RSLPLGLIGHKRSVYWHVIGMTPEVHSIFLEGHTFLVRNHRQASLEISPTFLTAOTL 318  
Db 240 GTMDPDTVCAHDIHNLHMGMSGPELFIHNGQVLEQNHKHSVITLVASTSTTANMT 299  
Qy 319 LMDLGOFLFLFCHTISHQNDHMEAYKVDSCPREPDRKMKNNDEADYDDDLTDEMDYVR 378  
Db 300 VGEGRKVIISLTPKHIQAGQAYVADIKKCPKTRKLKTPRE----- 342  
Qy 379 FDDDNPSFPIQIRSVAKKPKTMVNHIAEEDDMYARVLPDRRSYKSOYLANGPQRI 438  
Db 343 -----ORHMKRWEFTIAAEVIMDAPVAPNMMKRYRSOHLNFSMOI 387









C:Function:  
A:Description: catalyzes the oxidation of free iron(II) to iron(III) coupled with the re  
A:Note: iron(III), but not iron(II), is the form bound and transported by transferrin  
C:Superfamily: ferroxidase; ferroxidase repeat homology  
C:Keywords: acute phase; alternative splicing; copper; duplication; glycoprotein; oxidat  
F:1-19/Domain: signal sequence #status predicted <Sig>  
F:20-1069/Product: ferroxidase, long form #status predicted <MATL>  
F:20-1060/Product: ferroxidase, short form #status experimental <MANS>  
F:20-499/Product: ferroxidase 67k chain #status experimental <K67>  
F:23-357/Domain: ferroxidase repeat homology <FO1>  
F:373-718/Domain: ferroxidase repeat homology <FO2>  
F:501-905/Product: ferroxidase 50k chain #status experimental <K50>  
F:733-1059/Domain: ferroxidase repeat homology <FO3>  
F:907-1065/Product: ferroxidase 19k chain #status experimental <K19>  
F:138-200,276-357,534-560,657-718,874-900/Disulfide bonds: #status predicted  
F:127,388,526/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:293,338,343/Binding site: copper (His, Cys, His) (type 1) #status absent  
F:358/Binding site: carbohydrate (Asn) (covalent) (partial) #status predicted  
F:656,699,704,709/Binding site: copper (His, Cys, His, Met) (type 1) #status predicted  
F:994,1040,1045,1050/Binding site: copper (His, Cys, His, Met) (type 1) #status predicted

Query Match 22.08; Score 1729; DB 1; Length 1069;  
Best Local Similarity 32.1%; Pred. No. 2.4e-109;  
Matches 389; Conservative 197; Mismatches 419; Indels 208; Gaps 23;

Qy 5 LSTCFCLLRFCSATRRYLGLAVELSDMYQSDLG- -LPVDARPPRPVPSFPNT 61  
D 6 LGIFLFLCSTP-AWAKEHHYYIGIETTWDY-ASDHGEKKLISVTEHSNITLQNGPDRI 63  
Qy 62 SVYKKTFLFEVETVHLENIARPRPMGLLPTIOAEYVDYVITLKKMASHPVSLHVG 121  
D 64 GRLLKALYLDTTDETTFTTEKRPWGLPLGLITKAETGDUYVHLKMLASRPYFESHG 123  
Qy 122 VSYKASGAEYDDQTSQREKEDKVPFGSGHTYVQVLKENGPMASDPLCLTYSLSHV 181  
D 124 IITYKEHGAIPYDNTTDFQADKQVYRGEOYTYMLATEQSPGEGNCVTRHYSHI 183  
Qy 182 DLVKDLSGLIGALLVCRESGLAKEKTOFL-HKFTLLFAVDEGSKWSE- - -TKNS 234  
D 184 DAPKDASGLGLPLICKKSDLSDEKEKHIDREFVMSVVDENSWYLEINIKTYCSEP 243  
Qy 235 LMORBDASAPAKPMHMYGVNRSLSGLIGCHKRSYVMYVIGTGPPEVHSIFLEHT 294  
D 244 EKVDKDNEDFESNRKMSVNGITGSLGLSCMAEDRYKMTLFGMGNVDYHAAPFHQA 303  
Qy 295 FLVNRHQASLEISPIFTFLAQTLLMDGQFLFCHISSHQHDMGMAVYKDCPEEPOL 354  
D 304 LTNKMYRIDTINLPPATLFDAYVAQNPGEMWLSQNLNHLKAGLQAFQYQEC- - - 357  
Qy 355 RMRKNEAEYDDDLTDESEMYVRFDDNSPFIQIRSVAKKHKTWVHTYAAEEDMDY 414  
D 358 - - -KKSSKSD- - - - -NIRKHNVRH- - - - -YUAAEETILNMY 385  
Qy 415 APL- - - - -VLAPDRSYKSOVLNNGPQIRGRKYKVPFMAVTDFT- - -KTREAIO 462  
D 386 APSGIDIFTKNTLAPRGS- - -AVFEQGTTRIGSGYKLVREYTDLSFTNKRKEGEE 443  
Qy 463 HESGILGLLXGEVDTLLIFKQNASRPYNIYPHG- - - - -TDVRLPY- - -SRRLP 511  
D 444 EHLGLGVIAVEGDTIRTVFNHKGAVPLSIEPLGVNFKNNGESTYSPNYNQSRSPV 503  
Qy 512 KGVNHLKDFLLPGLGEIRKYKTVVEGDPYKSDRCLTRYSSPYNMERDLASGLIGPL 571  
D 504 PSASH- - - - -AAPTEFTYEVTVKEVGPITNADPCLAKMYSAVDPKIDFTGLIGPMK 558  
Qy 572 ICYKESYDQROQNIIMSDRNVILFSVDENRSWYLTENIOFLPNPAGVQLEDEPQSN 631  
D 559 ICKGSLHANKROKDYKEFLFPVDEPNESLLEDNIIRFTTAPRDVQKDEDEPQSN 618  
Qy 632 IMHSINGVFDSDLO-LSVGLHEVAWYILSIGAOTDLSVFEFSGVTFKHKMYEDTLTLF 690

Qy 619 KHSNNGMYGNQPLTMCCKGDSVYVLYFSAGNENADVIGIYSGNTLYMRGERDTANLF 678  
D 651 PFGSGTVMSENPGLMILGLCHNSDFRNKGMTALLKVSCKDNQDYEDYEDISAYLL 750  
Qy 679 POTSLLTMMPPESTFVNECLTLDHYGKOKKYTVNOCROS- - -EDS- - - 725  
Qy 751 SKNNMIEPRSFQSNRHSSTQKQFNAPRYLKRQRELTRTLOSDEEIDYDTISVE 810  
D 726 - - - - - 725  
Qy 811 MKKEDFIYDEDENSPSPFOKTRHYFAVERLMDYGMSSP- - - - -HVLNRNQ 862  
D 726 - - - - -TFYLGENTYIAAVEVMD- - -SPQREKEKELHHLQEQNV 763  
Qy 863 SGSV- - - - -POFKVYVQEFIDSGTQPLVYGLLENHILGLPYITRAVEDNINVT 913  
D 764 SNAFLDKGEFYIGSKYKVVYQYDSTFRVPERKAEENHILGIPOLHADVGDKVKII 823  
Qy 914 FRNQSAPYSPFSSLSIYEDQROGAEPRKNPVK- - -PNETKYFWKVOHNHAPTKDEPD 970  
D 824 FKNMATRYSIHA- - - - -HGVTESSTVYTPTLDEGTLTYWKIPRNSGACTEDSA 873  
Qy 971 CKAAVFSVDVLEKDVHSGILGPLVCHTNTLPNAGROVYQEPALFETTFDETKSWYF 1030  
D 874 CIPMAVYSTVDQKDYLSGLIGPLVCRRPYLKVNPRRKL- - -EFALLFVDEDNESWYL 931  
Qy 1031 TENMERNCRAPCNOMEPTTKENTRFHAINCYIMDTLPGLYMAODQIRMYLMSGSNE 1090  
D 932 DDNITKDYDHPEKVKKDEEPEESNKMHAJNRMFGNLQGLIMHAGDEVDENYLMCMGNEI 991  
Qy 1091 NISHIFSGHVFTVKKREYKMAIYNLPGVEFPEMLPSKAGIYRVCLIGENHAGMS 1150  
D 992 DLHTVHFHSHQYKHKRVYSSDVDFIRPGYTQLEMPRTPGIWLHLCHVTDRHAGAE 1051  
Qy 1151 TLFLVYSNKCQTP 1163  
D 1052 TTYVYLQNEGEVP 1064

RESULT 8  
A:35210  
A:ferroxidase (EC 1.16.3.1) precursor - rat  
N:Alternate names: ceruloplasmin  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 10-Sep-1999  
C:Accession: A35210; A41753; A29564; S21692  
R:Fleming, R.E.; Gitlin, J.D.  
J. Biol. Chem. 265, 7701-7707, 1990  
A:Title: Primary structure of rat ceruloplasmin and analysis of tissue-specific gene  
A:Reference number: A35210; MUID:90237081  
A:Accession: A35210  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-1059 <FE>  
A:Cross-references: GB:J05424  
R:Fleming, R.E.; Gitlin, J.D.  
J. Biol. Chem. 267, 479-486, 1992  
A:Title: Structural and functional analysis of the 5'-flanking region of the rat ceru  
A:Reference number: A41753; MUID:92112697  
A:Accession: A41753  
A:Molecule type: DNA  
A:Residues: 1-48 <FL2>  
A:Note: the authors translated the codon GAA for residue 40 as Gly, GAA for residue 4  
R:Aldred, A.R.; Grimes, A.; Schreiber, G.; Mercer, J.F.B.  
J. Biol. Chem. 262, 2875-2878, 1987  
A:Title: Rat ceruloplasmin. Molecular cloning and gene expression in liver, choroid p  
A:Reference number: A29564; MUID:87137545  
A:Accession: A29564  
A:Molecule type: mRNA  
A:Residues: NSG, 215-216, 'Y', 218, 'PAT', 222, 'F', 224-226, 'E', 228, 'LL', 231, 'D', 233-235,  
A:Experimental source: liver  
A:Note: the authors translated the codon GCG for residue 60 as Gly and GTG for residu



R: Ryan, T.P.; Grover, T.A.; Aust, S.D.  
 Arch. Biochem. Biophys. 293, 1-8, 1992  
 A: Title: Rat ceruloplasmin: resistance to proteolysis and kinetic comparison with human  
 A: Reference number: S21692; MUID: 92117681  
 A: Accession: S21692

A: Molecule type: Protein  
 A: Residues: 20-29, 'Q'; 902-910 <RYA>  
 C: Superfamily: ferroxidase; ferroxidase repeat homology  
 C: Keywords: copper; glycoprotein; oxidoreductase; plasma  
 F: 1-19/Domain: signal sequence #status predicted <SIG>  
 F: 20-109/Domain: signal sequence #status predicted <SIG>  
 F: 22-35/Domain: ferroxidase repeat homology <FO1>  
 F: 372-712/Domain: ferroxidase repeat homology <FO2>  
 F: 727-1053/Domain: ferroxidase repeat homology <FO3>

Query Match 21.4%; Score 1681; DB 1; Length 1059;  
 Best Local Similarity 32.0%; Pred. No. 4,3e-106;  
 Matches 385; Conservative 188; Mismatches 429; Indels 200; Gaps 23;

5 LSTCFLLRRCFSATRRYVLAVALSWDYMQ-SDLGEL-PVDARFPPEVPKSPFPTS 62  
 6 LSAFLFL-HSLAMTRREKHYIGITEAVWDVYASGSEKELISVDTEQSNFYLRNGPDRIG 64  
 63 VVYKTLVEFTVHLFNIAKPRPMGLIGPTIOAEVYDTVYITLKNMASHVYLHAGV 122  
 65 RYKKAALSYEDYDGFETFTIDKPAWLGELGVYKAEVGDVSVHYVKNFASRYTFHAHV 124  
 123 SYWRASEGAEDDQTSOREKEEDKVPGGSHYVMOVLKENGPMASDPLCTFYSTYSHV 182  
 125 TYTNANEGALTFPDNTDTPORADCKLPPGOOYLVLRA-NEPSGEGDSNCVIRITHSHV 183  
 183 LVKDLNSGLIGALLVCRGSLAKERTQTL-HKFTLLFAVDEGKSWHSEKNSIM----- 236  
 184 APKDIASLIGLICKKSLKKEENIDGELVMSVDELSMYLDNKTFCSEPE 243  
 237 -QDRDASARAPKMHYVNGYVNSLPGICGHRKSVYHVGMTTPVHSTFLEGHTF 295  
 244 KVDKNEDEFOESNRMYSINGYTFGSLPGSMCAEDRVKVKYLTGCMGEVNVHSLFPGOAL 303  
 296 LVNRHROASLESPTPLTFLAOTFLMDLGOFLFCHHSQHOGMEVYVVDSCPEPQAR 355  
 304 TSKNHTDITINLPATLLDVSAVQAOPGVWMLSCONLNLKAGLQAFVORDC----- 356  
 356 MKNNEADYDDDLTDSEMDVYRFDDNSPSFQIRSVAKHKPKTWVHYIAAEEEDMDYA 415  
 357 ---NKPSP---DDIDDRHY-----RH-----YIAAETITMDVA 385  
 416 P-----LVLPDRDSKSOYLNNGPORIGKRYKVKVPMATYDTET---KTRERAI 461  
 386 PSCGDTFTGENTLSGSDSRVFEQ---GATRIGSYKKLYREYTDSTFNKRKRGCPD 441  
 462 QHESGIIAPLLYGEVDPDLLIFKNOASRPVNIYPHGITDVR---PLDSRRLPKVKVHL 517  
 442 EHLGILCPVIAEYGDIIIRYFNHKGOPFLSIOPMGVFTFENETYYG---PDGRSSK 498  
 499 QASHVAPKETFTYEWTVPEKMGPTADPYCLSKMYSGVDLKDITGLTGKIPKICKKGS 558  
 518 KDFLLPEIRKYKWTYVVEDGPTKSDPRCLTRYSSFFVNMERDLASGLIGLLCYKES 577  
 578 VDQGNQIMSKRNVILSVFDEKNSWYLTENIORPLRPAGVQLEDPFQASNIHMSIN 637  
 559 LLAOROKDVAKKEFLFATFEDENESLLDDNIRMETTPRENVDKDEDEFOESKNHNSN 618  
 638 GYVFDSDLG-LSVCLHEVAYWYTLIGAOTDPLSVFSGYTFKHKMYEDTLTLFPSSGFT 696  
 619 GFMKGNLGLNLCGEISYVWYLFYSAGNEADVHGYTFSGTYVLSKGRDPTANLFPKSLT 678  
 697 VFMSENENGLWILGCHNSDFRNGMTALLKVSACDKNTGDYEDSEYDYSAYLLSKNNAI 756  
 679 LLMFPDTEGSPFVBECLTTHYTGKMKOKYTVNOC-----KQGEEDTVLY----- 722  
 \*QY 757 EPRSPFONSRRHPSTROKOPNATPVYKRRHQREITRTTLOSDEEDIDYDTISVEMKEDF 816

Db 723 ----- 722  
 QY 817 DIVDEENQSPRSFOKTRHYFLAVERLMDYGMSSPHVLNRRAGSV----- 866  
 Db 723 -----OGERTYIAAVEVMDYSPSRDWMELHHLQEDQVNSAFIDKEEF 767  
 QY 867 ---PQKRVYFOEFTDGSFTOPLYRGELNHLGLAPYIRAEVEDNIWTERFNSAPRS 923  
 Db 768 FLSKSKVYVREFDTSTREQYKRAAEEHGIIGLPLIHADVGAKKVVFKNMATRYS 827  
 QY 924 FYSSLISYEDOROGAPPKKNFYK---PNETITYKVKOVHNAAPKDEPDCAKAYESDV 980  
 Db 828 IHA-----HGKTRKSTVAPLTPGEVRYTYIQIIPERSAGTEDSPCTPMAYYSTV 877  
 QY 981 DLEKDVHSGLIGPLVLC---HTNTLNPARGROVTVOEFALFTFIDEKSWFTEFMRN 1037  
 Db 878 DRKDKLISGLIGPLVLCRSYKVENPK-----KMERISLFLVDEENSWYLDNINITY 932  
 QY 1038 CRAPCNIDEDPTFKENYRFAHNGYIMDTPLGLVMAODORIRWYLLSGSNENIHSIH 1097  
 Db 933 PDHPEKVNKDNEDEFISNKMHAINGKMGNLQGLTFHWVGDVWYVMAGNEIDLHTVHF 992  
 QY 1098 SCHVFTVRKKEEKYKALVLYPGVEFTVEMLPKSKACIMRVECLIEHLHAGSTLELYVS 1157  
 Db 993 HGHSFQYKRGHSSDVFDFEFTYQTLFEMPOTGTWLLHCHVYDHIHAGVWYTYVLP 1052  
 QY 1158 NK 1159  
 Db 1053 NO 1054

RESULT 9  
 A44258  
 factor VIII-associated gene B hypothetical protein - human  
 C: Species: Homo sapiens (man)  
 C: Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 18-Jun-1999  
 C: Accession: A44258  
 R: Levinson, B.; Kenwick, S.; Gamel, P.; Fisher, K.; Gitschler, J.  
 Genomics 14, 585-589, 1992  
 A: Title: Evidence for a third transcript from the human factor VIII gene.  
 A: Reference number: A44258; MUID: 93052386  
 A: Accession: A44258  
 A: Status: preliminary  
 A: Molecule type: mRNA  
 A: Residues: 1-216 <LEV>  
 C: Cross-references: GB:M90707; MUI:9182316; PIDN:AA58466.1; PID:9182317  
 F: 1-53/Domain: discolidin I amino-terminal homology #status atypical <DN1>  
 F: 57-210/Domain: discolidin I amino-terminal homology <DN2>

Query Match 14.0%; Score 1104; DB 2; Length 216;  
 Best Local Similarity 100.0%; Pred. No. 6,6e-66;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1264 VFGNVDSGKIRNINFPPIIARYIRLAPTHYSIRSTYLMELMGCDLNSCNPGLMESKA 1323  
 Db 9 VFGNVDSGKIRNINFPPIIARYIRLAPTHYSIRSTYLMELMGCDLNSCNPGLMESKA 68  
 QY 1324 ISDAQITASSYFTNMFATYSPSKARLHLOGRSNANRPVYNNKREKLYQDFQTKMYTVYT 1383  
 Db 69 ISDAQITASSYFTNMFATYSPSKARLHLOGRSNANRPVYNNKREKLYQDFQTKMYTVYT 128  
 QY 1384 TGVKSLTSMVYKELTSSSDGHWMTLFPONGVYKVGQODSFTPVVNSLDPPLTR 1443  
 Db 129 TGVKSLTSMVYKELTSSSDGHWMTLFPONGVYKVGQODSFTPVVNSLDPPLTR 188  
 QY 1444 YLRHPQSVVHQAIALMEVYLGCEADLY 1471  
 Db 189 YLRHPQSVVHQAIALMEVYLGCEADLY 216

RESULT 10  
JC4915  
ags protein precursor - rat  
N:Alternate names: O-acetyl-Gd3 ganglioside.  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 26-Sep-1996 #sequence\_revision 01-Nov-1996 #text\_change 20-Jun-2000  
C:Accession: JC4915  
R:Ogura, K.; Nara, K.; Matanabe, Y.; Kohno, K.; Tai, T.; Sanai, Y.  
Biochem. Biophys. Res. Commun. 225, 932-938, 1996  
A:Title: Cloning and expression of cDNA for O-acetylation of GD3 ganglioside.  
A:Reference number: JC4915; MUID:96374422  
A:Accession: JC4915  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-427 <OGN>  
A:Cross-references: DDBJ:D84068; NID:g1620006; PIDN:BA412210.1; PID:g1620007  
A:Experimental source: CST cell  
C:Comment: This protein is required for the O-acetylation of disialoganglioside sialic acid  
C:Genetics:  
A:Gene: ags  
C:Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolo  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:28-60/Domain: EGF homology <EG1>  
F:68-107/Domain: EGF homology <EG2>  
F:110-267/Domain: discoidin I amino-terminal homology <DNI>  
F:271-427/Domain: discoidin I amino-terminal homology <DNI>

Query Match 8.4%; Score 663; DB 2; Length 427;  
Best Local Similarity 40.3%; Pred. No. 2,26-37;  
Matches 142; Conservative 61; Mismatches 123; Indels 26; Gaps 6;  
QY 1138 ECLIGEHHAHAGMSTFLVYSNK-----COTPLGMAHGIRDPOTASGOY-- 1182  
DB 78 KCLVETDTRG--DIFETYLQCPVYSGIHCELGCELTGLGALADQSISASVYMG 135  
QY 1183 -----GOMAPKLARLHYSGSINAM--STKEPFWIKVDLAPMIHIGIKTGAROKFSSLY 1236  
DB 136 FMGIQRWCPRLARLHYSGSINAM--STKEPFWIKVDLAPMIHIGIKTGAROKFSSLY 1236  
QY 1237 ISOFITMYSIDGKKWQYRGSTGLMVFEGNVDSGIGIKHIFNPPIIARYIRLPHYYS 1296  
DB 196 LKTFKVAVSTDGRRFEEFQDESGTGDEKFEFGNODNNSLKINMNPDLTAQYIRLPHVSC 255  
QY 1297 IRLRLMELMCDLNSGSMRPGMESKASDQITASSFT--NMFA-TWSSKARLHOG 1353  
DB 256 RGCTLRRELLCELRHGESEPLGKNNTPDSQITASSSYKTNMRAFGWYHPLGRLDNOG 315  
QY 1354 RSNAMRPQVNNPKEMLOVDFQKTMKVTGVTQGVKSLTSMYKVEFLISSQDGHQWTLF 1413  
DB 316 KINMWTQSNNAKEMLOVDLCTOKKVGIIITGARDFGHIOYVASYKAHSDDGQWTVY 375  
QY 1414 FONGKVFQGNODSFPYVNSLDPPLIRLPHPOSWHQIALRMEVLGC 1465  
DB 376 EEOGTSKVFQGNLNDNSHKKNIFKPFMARVRYVPLPSMHNRIPLRELLGC 427

RESULT 11  
A36479  
milk fat globule membrane protein - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A36479  
R:Stubbbs, J.D.; Lekutis, C.; Singer, K.L.; Bui, A.; Yuzukl, D.; Srihivasan, U.; Parry, C.  
Proc. Natl. Acad. Sci. U.S.A. 87, 8417-8421, 1990  
A:Title: cDNA cloning of a mouse mammary epithelial cell surface protein reveals the ex  
A:Reference number: A36479; MUID:91046008  
A:Accession: A36479  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-463 <STU>  
A:Cross-references: GB:M38337; NID:g199142; PIDN:AAA39534.1; PID:g199143  
C:Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolo

C:Keywords: membrane protein  
F:28-60/Domain: EGF homology <EG1>  
F:68-107/Domain: EGF homology <EG2>  
F:147-303/Domain: discoidin I amino-terminal homology <DNI>  
F:307-463/Domain: discoidin I amino-terminal homology <DNI>

Query Match 8.3%; Score 657; DB 1; Length 463;  
Best Local Similarity 42.2%; Pred. No. 6,56-37;  
Matches 135; Conservative 63; Mismatches 110; Indels 12; Gaps 5;  
QY 1157 SNKCTPLGMAHGIRDPOTASGOY-----GOMAPKLARLHYSGSINAM--STKEPFS 1208  
DB 145 ASRCSTQMEGALADQSISASVYMGFMGIQRWCPRLARLHYSGSINAM--STKEPFS 204  
QY 1209 WIKVDLAPMIHIGIKTGAROKFSSLYISOFITMYSIDGKKWQYRGSTGLMVFEGN 1268  
DB 205 WIOVNLKRMKRVSGWMTQGSARAGRAEYLRKTFKVAVSTDGRRFEEFQDESGD-KEPFGN 263  
QY 1269 VDSGIGIKHIFNPPIIARYIRLPHYYSIRSLRMLGCDLNSGSMPLGMSKASDAQ 1328  
DB 264 LDNNSLKNNMNPDLTAQYIRLPHVSCIRGCTLRRELLCELRHGESEPLGKNNTPDSQ 323  
QY 1329 ITASSYFT--NMFA-TWSSKARLHOGRSNAMPQVNNPKEMLOVDFQKTMKVTGVTQ 1385  
DB 324 MSASSYKTNMRAFGWYHPLGRLDNOGKINAMWTQSNNAKEMLOVDLCTOKRQVGIITQ 383  
QY 1386 GVKSLTSMYKVEFLISSQDGHQWTLFONGKVFQGNODSFPYVNSLDPPLIRYL 1445  
DB 384 GARDGCHIOYVESYKVAHSDDGQWTVYEEOGSSKVFQGNLNDNSHKKNIFKPFMARVY 443  
QY 1446 RIHQSWHQAIALRMEVLGC 1465  
DB 444 RVLPSWMHNRITLRELLGC 463

RESULT 12  
T11743  
Pp47 protein - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C:Accession: T11743  
R:Enslin, M.; Vogel, T.; Calvele, J.J.; Thole, H.H.; Schmidke, J.; Matsuda, T.; Toe  
Bio. Reprod. 58, 1057-1064, 1998  
A:Title: Molecular cloning and characterization of P47, a novel boar sperm-associated  
A:Reference number: Z17325; MUID:98206817  
A:Accession: T11743  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-409 <ENS>  
A:Cross-references: EMBL:Y11683; NID:g2652927; PIDN:CAA72379.1; PID:g2652928  
A:Experimental source: testis  
C:Function:  
A:Description: may be involved in membrane remodeling and/or function as a zona pellu  
C:Superfamily: unassigned EGF-related proteins; EGF homology  
F:6-40/Domain: EGF homology <EGF>

Query Match 8.3%; Score 650; DB 2; Length 409;  
Best Local Similarity 39.3%; Pred. No. 1,66-36;  
Matches 139; Conservative 62; Mismatches 125; Indels 28; Gaps 6;  
QY 1138 ECLIGEHHAHAGMSTFLVYSNK-----COTPLGMAHGIRDPOTASGOY-- 1182  
DB 58 ECEVIDAHRG--DVFTFYICKCPHYGTHICETICNAPLMEGALADFOISASSMLIG 115  
QY 1183 -----GOMAPKLARLHYSGSINAM--STKEPFWIKVDLAPMIHIGIKTGAROKFSSLY 1236  
DB 116 FMGIQRWCPRLARLHYSGSINAM--STKEPFWIKVDLAPMIHIGIKTGAROKFSSLY 1236  
QY 1237 ISOFITMYSIDGKKWQYRGSTGLMVFEGNVDSGIGIKHIFNPPIIARYIRLPHYYS 1296  
DB 176 MKTFKVAVSTDGRRFEEFQGAESGDKIFMGNLNDNSGLKVMLEFVPLEVQYVRLVPIICH 235

[illegible][illegible]

A25945  
coagulation factor VIII - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 16-Aug-1988 #sequence\_revision 16-Aug-1988 #text\_change 21-Jul-2000  
C:Accession: A25945  
R:Toole, J.J.; Pittman, D.D.; Orr, E.C.; Murtha, P.; Wasley, L.C.; Kaufman, R.J.  
Proc. Natl. Acad. Sci. U.S.A. 83, 5939-5942, 1986  
A:Title: A large region (lapprox195 kba) of human factor VIII is dispensable for in vitro  
A:Reference number: A25945; MUID:86287369  
A:Accession: A25945  
A:Status: nucleic acid sequence not shown: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-869 <T00>  
C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase

|                       |                 |                   |          |            |
|-----------------------|-----------------|-------------------|----------|------------|
| Query Match           | 6.9%            | Score 543         | DB 2     | Length 869 |
| Best Local Similarity | 75.28%          | Pred. No. 1.1e-28 |          |            |
| Matches 100           | Conservative 14 | Mismatches 19     | Indels 0 | Gaps 0     |

[illegible]

Search completed: November 17, 2001, 13:23:23  
Job time: 72 sec

**THIS PAGE BLANK (USPTO)**